



## Review

# Electronics and data acquisition

Helmuth Spieler

Physics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

## ARTICLE INFO

Available online 5 May 2011

### Keywords:

Radiation detectors  
Readout systems  
Electronics  
Low noise analog electronics  
Signal processing  
Radiation resistance  
Power

## ABSTRACT

High energy physics detectors span a wide range of applications with greatly differing requirements. Although the detector configurations are very different, the application of only a few basic signal acquisition principles is required. The LHC required novel designs, but built on a wide range of previous developments that had been completed for other experiments. The high luminosity drove up the event rates, but multiple interactions per bunch crossing also made occupancy a major challenge. The large scale of detector subsystems imposed efficient designs where cost was a major consideration, but the difficulty of accessing detector components added reliability to the list of more severe requirements. Radiation damage, especially in the inner detectors, added additional crucial constraints. This paper will discuss electronics requirements, the configurations of major LHC detectors, and the readout systems. After a discussion of front-end implementations and radiation effects, systems with extreme performance requirements are described in more detail, i.e. silicon strip and pixel systems.

© 2011 Elsevier B.V. All rights reserved.

## Contents

1. Introduction . . . . .	198
2. Commonality of different applications . . . . .	198
3. Pixel detectors . . . . .	199
4. Leading edge developments – detector requirements at the LHC . . . . .	200
4.1. Event rates and occupancy . . . . .	200
4.2. Signals and noise . . . . .	201
4.3. Fast data readout . . . . .	201
4.4. Radiation resistance . . . . .	201
4.5. Technology, reliability, and cost . . . . .	202
5. LHC Detector configurations and subdetector signal sources. . . . .	203
5.1. ATLAS . . . . .	203
5.2. CMS . . . . .	204
5.3. LHCb . . . . .	205
5.4. ALICE . . . . .	207
6. Front-end configurations . . . . .	208
6.1. Performance specifications . . . . .	208
6.1.1. Electronic noise . . . . .	208
6.1.2. Input impedance . . . . .	209
6.2. Charge-sensitive front-ends and shapers . . . . .	210
6.3. Low-power front-ends . . . . .	210
6.4. Current sensitive front-ends (ATLAS liquid Ar calorimeter) . . . . .	211
7. Radiation resistance . . . . .	212
7.1. Sensors . . . . .	212
7.2. Electronics . . . . .	212
8. Examples with extreme requirements I: silicon trackers . . . . .	212
8.1. Coping with high rates . . . . .	213
8.2. Radiation damage . . . . .	213
8.3. Layout . . . . .	213

E-mail address: [helmuth.spieler@gmail.com](mailto:helmuth.spieler@gmail.com)

8.4.	Readout electronics	214
8.4.1.	CMS readout electronics	214
8.4.2.	ATLAS readout electronics	215
8.4.3.	Required signal-to-noise ratio in a binary readout system	215
8.4.4.	ATLAS SCT readout implementation	216
8.5.	Detector modules	217
9.	Examples with extreme requirements II: Silicon pixel detectors	218
9.1.	ATLAS pixel detector	218
10.	Conclusions	222
	Acknowledgment	222
	References	222

## 1. Introduction

High energy physics detectors span a wide range of applications with greatly differing requirements. Although the detector configurations are very different, the application of only a few basic signal acquisition principles is required. There are some basic components common to the readout systems:

- Front-ends, which include preamplifiers, pulse shapers, buffer memories, and digitizers.
- Preprocessors and drivers, which include zero-suppression, data configuration, and output drivers.
- Off-detector data processing.

The first two items are often located in or near the active detector region, where space, power, and material are major considerations, whereas the off-detector data processing is often far removed and thus offers more flexibility because space and power constraints are not as critical.

## 2. Commonality of different applications

Although the specific parameters can vary significantly, it simply means that the basic signal acquisition techniques must be applied accordingly. In today's high energy physics systems practically all detectors require conversion of the primary detection to electrical signals. There are exceptions in particle physics, e.g. bubble detectors in dark matter searches, but even these use electronic imaging to analyze the signals.

There are two basic types of detectors, direct and indirect signal production. In direct detection the absorbed energy is converted directly into charge, i.e. electron–ion or electron–hole pairs. The conversion configurations span a wide range:

- Drift chambers;
- Multi-wire proportional sensors;
- GEMs;
- MicroMegas;
- Liquid noble gas ionization chambers (calorimeters);
- Semiconductor detectors, e.g. strip and pixel detectors.

Indirect detectors are primarily scintillation detectors, where scintillation light is converted to an electrical signal through a variety of techniques:

- Photomultipliers with high gain;
- Hybrid phototubes that have less gain, but are also less sensitive to magnetic fields;
- Semiconductor photodiodes;
- Avalanche photodiodes (APDs);

- “Silicon photomultipliers”;
- Electronic luminescence.

The goals of energy measurement, position sensing, timing, mere hit detection, or some combination determine the required electronic sensitivity, response time, and dynamic range. The primary signal can be measured in terms of voltage, current, or integrated charge, all of which are related.

The absorbed energy  $E$  is converted into a signal charge:

$$Q_S = \frac{E}{E_{SQ}} \quad (1)$$

where  $E_{SQ}$  is the energy required to create a signal quantum, e.g. an electron–ion pair or electron–hole pair. The instantaneous signal is a current, whose magnitude for a given charge depends on the drift time of the signal charges within the detector, but when deposited on a capacitance  $C$  the signal charge translates into a voltage:

$$V = \frac{Q_S}{C} \quad (2)$$

where  $C$  is the total capacitance onto which the signal charge is deposited. In a practical system this is the sensor capacitance, the input capacitance of the amplifier, capacitances of connecting cables and additional circuitry, and stray capacitances. The full value of this voltage is only obtained when the input time constant, i.e. the product of the input resistance  $R_i$  of the amplifier times the total input capacitance  $C$  is much larger than the charge collection time. For high rates this time constant may be constrained to smaller values, so then the peak voltage is less. For a given electronic noise level this will degrade the signal-to-noise ratio. This also shows that the signal-to-noise ratio degrades with increasing capacitance  $C$ .

If energy measurement is the prime goal and the shape of the primary detector pulses varies, the best technique to measure the energy is to use an active integrator, called a charge-sensitive amplifier. As discussed below, these amplifiers provide a low impedance that transfers the signal current to a feedback capacitor in the amplifier circuit. The signal charge then converts to the amplifier output voltage, whose dependence on the total input capacitance  $C$  can be very small. In this case the electronic noise depends on the capacitance  $C$ , so the signal-to-noise ratio still degrades with increasing capacitance.

The degradation of signal-to-noise ratio with increasing detector capacitance is a general result for many systems, but not for the same reason, as noted above for charge-sensitive amplifiers (for a simple explanation see Ref. [1]).

The signal current can also be sent through a current-sensitive amplifier, i.e. an amplifier where the input time constant  $R_i C$  is much smaller than the charge collection time, and then integrated

either in analog circuitry or by digitizing and then applying digital processing. In an ideal current-sensitive amplifier the noise can be independent of capacitance, but in reality the effect of parasitic noise sources can depend on detector capacitance, so it should always be considered.

The same principles also apply to detectors with internal gain, e.g. gas proportional chambers, GEMs, MicroMegas, APDs, etc., but now the signal is larger, so electronic noise is less important. Here the statistical fluctuations of the gain process often determine the energy resolution.

In scintillation detectors the scintillation light is also converted to an electronic signal. In a photomultiplier tube the primary output signal is a current, often simply sent through a resistor whose voltage is measured. Integrating the output current pulse yields a charge that is proportional to the energy deposited in the scintillator, so the basic techniques described above also apply, as also for other photodetectors. Again, with high gain prior to feeding the electronics, the electronic noise level is less important, but measurement of the signal magnitude or its timing may be the priorities. Overall, energy resolution depends on the ultimate signal-to-noise ratio, where the noise is the result of statistical fluctuations either in the detector, the electronics (i.e. electronic noise), or both.

In timing measurements signal-to-noise ratio is also important, but in addition the time dependence is key, i.e. the slope-to-noise ratio. For a time variant signal  $S(t)$  the time resolution:

$$\sigma_t = \frac{\sigma_n}{dS/dt} \quad (3)$$

where  $\sigma_n$  is the noise level, again the result of statistical fluctuations either in the detector, the electronics, or both. Since reducing the response time of the electronics, i.e. increasing  $dS/dt$ , requires an increase in frequency bandwidth, this increases electronic noise, so improving time and energy resolution require opposing electronics requirements. Both detector and electronics contributions are discussed in Ref. [1]. Although Ref. [1] targets semiconductor detector systems, the basic principles in optimizing front-end electronics apply to all detectors, although the combination of design parameters is quite different.

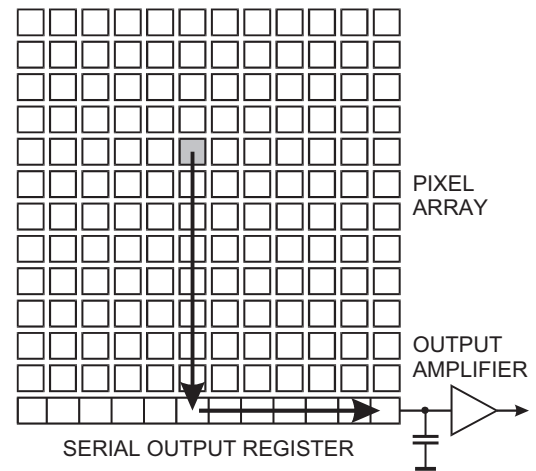
Detector designers commonly follow recipes that have worked in past designs. However, changes in the detector environment can lead to different paths and these must be recognized to determine which way to go. When the SSC was proposed the high luminosity of  $10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  led the recognized experts to claim that tracking detectors were impractical. As time has shown, this was a conclusion based on a narrow perspective and the lack of a broad understanding of physics. Developing novel detectors for high energy physics requires the understanding of a broad range of physics and the ability to recognize which directions to pursue and what is technologically practical.

### 3. Pixel detectors

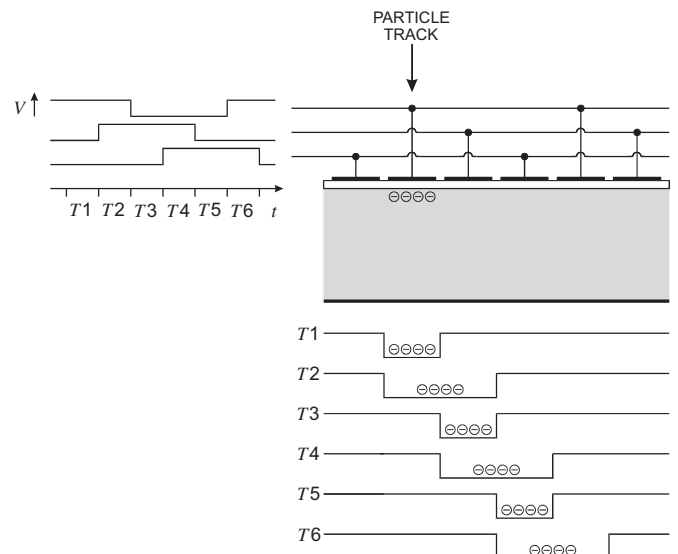
Pixel detectors have become increasingly important in high energy physics. In the LHC they are essential to separate tracks near the interaction region. Future tracking systems for the ILC require high single-track resolution to resolve the components of jets, so highly segmented devices with pixel sizes of about  $20 \mu\text{m}$  are needed, much smaller than used at the LHC. A major difference with respect to detectors at hadron colliders is that full coverage extending to small forward angles is crucial. In current detectors material tends to increase substantially at forward angles because that is where cabling and cooling lines are brought out (see Section 5). The ILC or similar future systems will require different concepts.

The first pixel detectors were charge coupled devices (CCDs). Fig. 1 shows the basic readout principle. Signals from the individual pixels within a column are transferred pixel-to-pixel and then fed into a serial output register. The signal charge is transferred from one pixel to the next by changing the relative potentials, as shown in Fig. 2. This serial readout is performed for all pixels, so for a transfer rate of  $R$  MHz and  $N$  pixels, the readout time is  $N/R \mu\text{s}$ . Increasing the readout clock rate will reduce the readout time, but also increase the power dissipation. For example, in the VXD3 vertex detector [2] that was successfully used at the SLD, a readout rate of 200 kHz per pixel led to peak currents of 1.3 A, which also raises cross-talk issues. The readout time per frame has been commonly reduced by subdividing the pixel array into four quadrants and more recently by having multiple readouts that serve several columns [3] or including a readout per column [4].

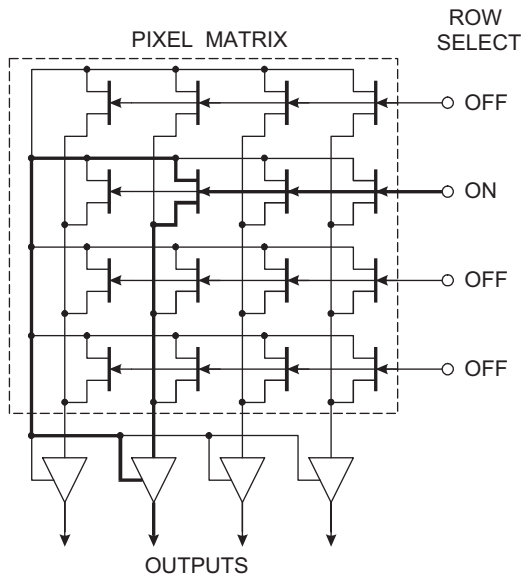
CCDs can be made with pixels of just a few microns size and with depletion depths in the hundreds of microns [5]. They can



**Fig. 1.** The pixels in CCDs are read out sequentially. The individual signals from all pixels in a given column are transferred to a serial output register, whose output is fed to the output amplifier. This is done sequentially for each column.



**Fig. 2.** The electric field applied to the semiconductor bulk moves the signal charge deposited in the bulk to the readout plane. The electrodes are capacitively coupled to the bulk, so changing the potentials on the individual electrodes will shift the charge from one electrode to the next. This figure shows how this is done in a three-phase clock pulse system.



**Fig. 3.** The DEPFET is read out by selecting a specific row, the second in this example. Then the corresponding pixels in all columns are available, which can all be read out or only the desired ones.

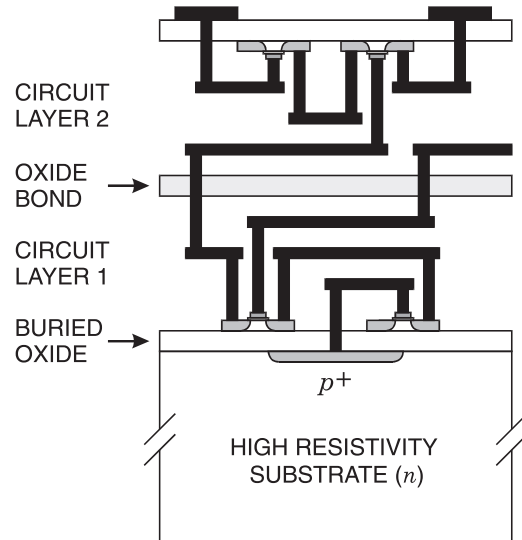
also be thinned to minimize material. An alternative is the DEPFET detector, which allows pixel selective readout [6], where the pixel is formed by a monolithically integrated transistor. Pixels are read out by selecting the corresponding  $x$  and  $y$  lines as shown in Fig. 3. However, in many applications it is necessary to include some circuitry on each pixel, e.g. to provide time stamping and local event storage. This can be achieved in monolithically integrated pixel arrays (MAPS) and hybrid pixel arrays.

MAPS use standard CMOS ICs where the epitaxial layer adjacent to the circuit plane is used as the active detector region. The field is not well controlled, but since the collection path is small, collection times are acceptable. Especially in modern “deep submicron” technologies this provides space for basic front-end circuitry [7]. MAPS are now also used in other fields, e.g. high-resolution dynamic electron microscopy [8]. The arrays are limited to the size of a chip, currently about  $3\text{ cm}^2$ , so multiple chips must be placed together with care to minimize dead space. A pixel size in the  $10\text{ s}$  of microns still limits the amount of circuitry that can be accommodated, but integrated circuit technologies that provide multiple circuit layers as illustrated in Fig. 4 will offer more possibilities [9,10]. These structures are often referred to as “SOI” or “3D”, not to be confused with other previously developed IC and detector structures that use the same names.

The most flexible pixel devices are hybrid arrays, key components in LHC detectors, where a large sensor array is used as the base onto which multiple chips are mounted. This also allows sensor materials other than silicon, e.g. diamond, to be used with a wide range of IC technologies. The pixel circuitry can include many functions, as illustrated in the ATLAS pixel detector that is described in more detail in Section 9.

#### 4. Leading edge developments – detector requirements at the LHC

The LHC presented a major challenge in developing practical detectors. The high luminosity imposes a large event rate with multiple interactions per bunch crossing and also makes radiation damage a major issue. Furthermore, the large size of detectors with the large number of channels requires a complex internal



**Fig. 4.** Multi-tier ICs combine multiple layers of electronics isolated by intermediate oxide layers. A high-resistivity substrate is utilized as the sensor. Sensor layers of  $350\text{ }\mu\text{m}$  thickness have been implemented. For fabrication steps see Refs. [9,11].

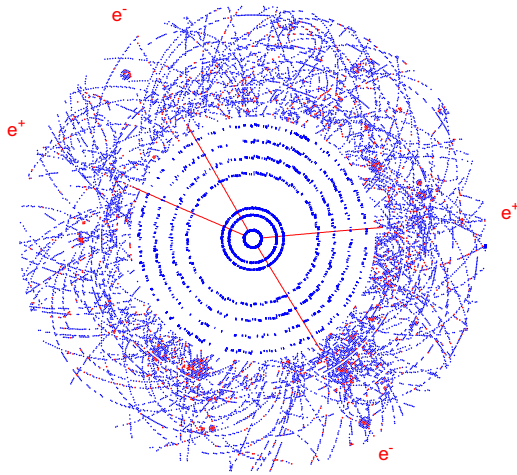
infrastructure. Unlike many previous detectors where individual subsystems could be assigned to individual institutions, the design and fabrication had to be distributed to groups spread all over the world. Budget constraints were also a major challenge. However, detectors are now in operation and are largely meeting the design requirements. Although this was in many respects a novel enterprise, it built on a wide range of previous developments, e.g. the Superconducting Supercollider (SSC), which for the same physics goals was planned for a higher energy (20 TeV center-of-mass) and 10% of the LHC luminosity, the CDF and DØ detectors at Fermilab, which gained extensive experience in dealing with the high occupancy compared to lepton colliders, and other experiments that developed complex detector and readout systems. Although many LHC designs are designated as novel, they utilize well-established principles, albeit sometimes in somewhat novel configurations. Many physicists are highly specialized, so they had to learn how to adopt techniques from other subsystems.

The following sections cannot cover all aspects of the readout systems in the LHC detectors. It will emphasize the extreme requirements imposed by the LHC’s high luminosity and pick individual examples from ATLAS, CMS, LHCb, and ALICE to illustrate the range of requirements and various solutions. In the last sections the silicon strip and pixel detectors in ATLAS and CMS will be discussed in more detail.

A general purpose detector includes vertexing, precision-tracking in a magnetic field (2 T in ATLAS, 4 T in CMS), calorimetry (electromagnetic plus hadronic), and muon detection. Fig. 5 shows an axial view of a simulated event.

##### 4.1. Event rates and occupancy

The LHC is designed for a particle bunch crossing rate of 40 MHz, i.e. 25 ns between bunches. At a luminosity of  $10^{34}\text{ cm}^{-2}\text{ s}^{-1}$  roughly 20 interactions will occur per bunch crossing with a total number of about 1000 tracks. Overall this leads to an event rate of about  $10^9\text{ s}^{-1}$ . Detectors must be capable of separating the individual interactions and assigning the multiple tracks accordingly. This also requires sufficient time resolution to separate events from different bunch crossings. It imposes requirements on tracking systems, but also on calorimeters, which must provide sufficient energy resolution together with



**Fig. 5.** Axial projection of a simulated Higgs to four electrons event at LHC design luminosity. Since the tracks are spread out along the beam axis, the  $z$ -resolution of the silicon detectors (inner rings) is essential for pattern recognition, although the high hit density of the outer straw chambers makes the tracks more apparent in this representation. (Figure from the ATLAS TDR [12]).

position and time information to associate data with the correct interactions.

The high hit rate can easily lead one to believe that very high speed electronics are needed. However, maintaining the required signal-to-noise ratio when reducing shaping times drives up the required power. Segmentation, on the other hand, is much more efficient, as subdividing the detector into many small elements reduces the hit rate per element. For example, at  $r_{\perp} = 30$  cm the hit rate on a strip electrode of  $50\text{ }\mu\text{m}$  width and  $10$  cm length, i.e. an area of  $5 \times 10^{-2}\text{ cm}^2$ , is about  $10^5\text{ s}^{-1}$ . This corresponds to an average time between hits of  $10\text{ }\mu\text{s}$ , so longer shaping times than implied by the overall event rate are allowable, which translates to lower power for a given noise level. As discussed in Section 6, with careful design the power requirements do not increase in highly segmented detectors (for a more detailed discussion see Ref. [1]).

An additional problem is the large number of events per crossing. Since the interaction region is spread out, vertex reconstruction with the appropriate  $z$ -resolution can resolve individual interactions. Again, segmentation helps. If a detector element is sufficiently small, the probability of two tracks striking it within one crossing can be negligible. Tracks from different crossings can be separated if the electronics are capable of single-bunch time resolution, that is  $25\text{ ns}$ , and time-stamped data are stored.

The spatial density of hits is especially high at small radii. Semiconductor detectors are well-matched to these requirements and they are key components in all LHC experiments. Patterning the sensor at the “ $\mu\text{m}$ -scale” is straightforward and monolithically integrated electronics can be mounted locally with on-chip multiplexing or sparsification to reduce the cable plant. The fast collection times of semiconductor are very advantageous; achieving collection times  $< 25\text{ ns}$  with a  $300\text{ }\mu\text{m}$  thick sensor is quite practical.

#### 4.2. Signals and noise

The time resolution requirements drive front-end electronics toward short shaping times. This increases the voltage noise contribution, but – as explained below – reduces the sensitivity to increased current noise due to radiation damage. Unlike many previous detectors where shaping times were typically in the  $200\text{ ns}$  to  $1\text{ }\mu\text{s}$  range, the shaping times in LHC silicon tracking systems are in the tens of ns range. Signal levels can degrade with increasing radiation loads and in systems that are designed to

minimize power, the noise margins must be carefully chosen. Another potential noise contribution that must be carefully controlled is cross-talk from the readout to the sensor input. At  $e^+e^-$  colliders the time between bunches was typically so large that signal acquisition and data readout did not have to occur simultaneously, so “common mode noise” was tolerated and widely considered to be unavoidable. This was largely due to basic design flaws and could be held at negligible levels.

Digital signal processing has become a practical solution for many more applications because of the significant ongoing improvements in fast digitizers and digital processing circuitry. There are some detector subsystems in LHC where digital signal processing is appropriate and advantageous (e.g. see LHCb and ALICE), but this is not a general conclusion. Where integration of sensors and electronics in a small volume is required, both circuit area and power dissipation are crucial considerations. Furthermore, these are special purpose systems. The electronics are specifically tailored to the sensor and application and do not need to be modified during the course of the experiments (the inevitable upgrades notwithstanding). Furthermore, for many sensors simple analog filters provide results that are only slightly inferior to the optimum filters that a DSP system would allow, but they achieve this with much lower power.

#### 4.3. Fast data readout

The interaction rate does not allow all data to be read out, so trigger systems are used to apply filters that enhance the ratio of desired events to background events and reduce the readout rate to manageable levels. Power and material constraints limit the maximum readout rate, typically to about  $10^5\text{ s}^{-1}$ , so local event storage is required in many systems to then read the full data upon receipt of a trigger. Fig. 6 shows the ATLAS trigger system as an example. Overall throughput is bounded at both the input, i.e. the data transmission bandwidth from the detector, which is limited by power considerations, and the output, limited by the storage media. The intermediate rates are a compromise to match these two end conditions, which means that the efficiency for desired events can be severely compromised. Increasing the processing speed for complex events can substantially increase the yield, as has been implemented by CDF [13,14].

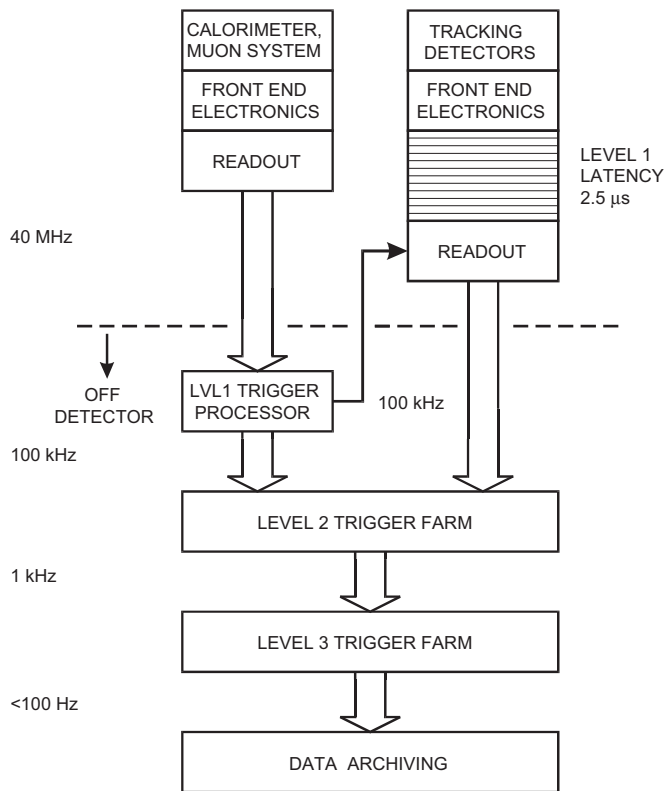
To reduce material in the data lines to the external readout systems most LHC detector modules feed optical converters (PIN/VCSEL arrays) and the data are sent serially at GHz rates by optical fibers.

#### 4.4. Radiation resistance

Radiation damage in both the sensors and the electronics will affect overall performance. There are two sources of particles, the interaction region and neutron albedo from the calorimeter. Estimated fluences per year at design luminosity expressed in equivalent  $1\text{ MeV}$  neutrons are  $5 \times 10^{13}\text{ cm}^{-2}$  at  $r=10\text{ cm}$  and  $2 \times 10^{13}\text{ cm}^{-2}$  at  $r=30\text{ cm}$ . The corresponding ionizing doses are  $30\text{ kGy}$  ( $3\text{ Mrad}$ ) at  $10\text{ cm}$  and  $4\text{ kGy}$  ( $400\text{ krad}$ ) at  $30\text{ cm}$ . In reality, complex maps are required of the radiation flux, which is dependent on local material distribution. Fig. 7 shows the hadron and neutron fluences over one year in the inner ATLAS detector. At small radii the two are about equal, but at larger radii the calorimeter albedo dominates.

A specially designed “BiCMOS” process that combined bipolar and CMOS devices was developed for LHC applications. Its radiation resistance proved to be adequate for most applications, but the fabrication yields were poor. This led to increased effort in testing ICs before fabricating detector modules. Pixel detectors in the inner radii required higher radiation resistance and initial





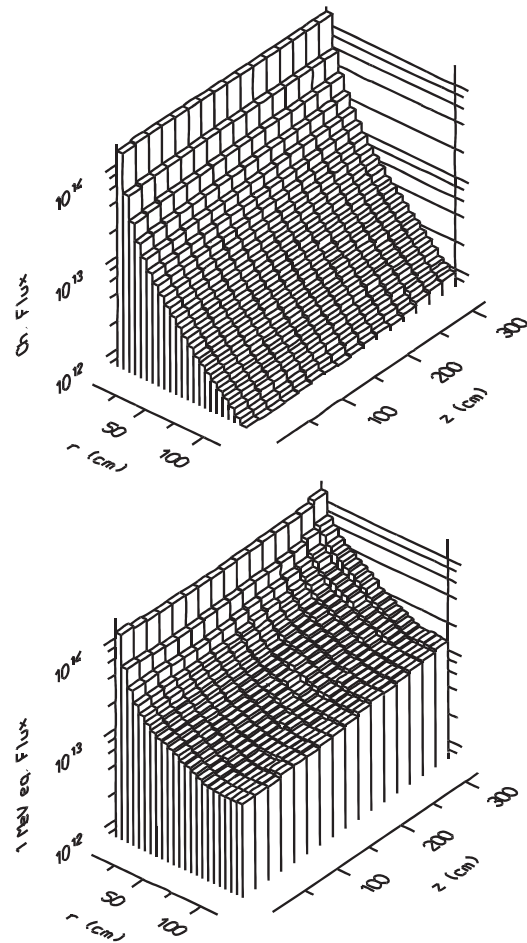
**Fig. 6.** ATLAS trigger configuration using calorimeter and muon detector data to derive the L1 trigger, which enables readout of the tracking detectors.

designs utilized special radiation-resistant fabrication processes. However, these became prohibitively expensive just before going into fabrication. Fortunately, developments in standard commercial “deep submicron” CMOS process led to much reduced radiation effects and circuits utilizing  $0.25\ \mu\text{m}$  feature size achieved much improved radiation resistance with excellent yields [31]. This was made possible by studies performed at CERN that followed up on the potential advantages of thin gate oxides that are inherent to the “deep submicron” processes. This also illustrates how basic investigations not linked to specific experiment projects can have a major effect on practical implementations. Immediate solutions are not always guaranteed, but complex detectors build on a wide range of parameters and experience, so basic detector research should be well supported by funding agencies.

In the course of developing silicon detector systems many unexpected problems arose and invariably solutions were found. Key is the use of a highly developed technology, which provides performance reserves and design flexibility. It is tempting to favor a new technology because it appears to offer a “silver bullet” against one specific problem, but systems depend on the interplay of many design considerations. For this reason many “advanced” technologies have fallen by the wayside. Often, existing technologies must simply be viewed from a different perspective with a thorough understanding of the underlying physics and its application, typically a major problem.

#### 4.5. Technology, reliability, and cost

Although high event rates and radiation damage appear to be the major challenges, the large size of the detector and the huge number of readout channels increase the amount of matter associated with the infrastructure, i.e. the cabling to provide

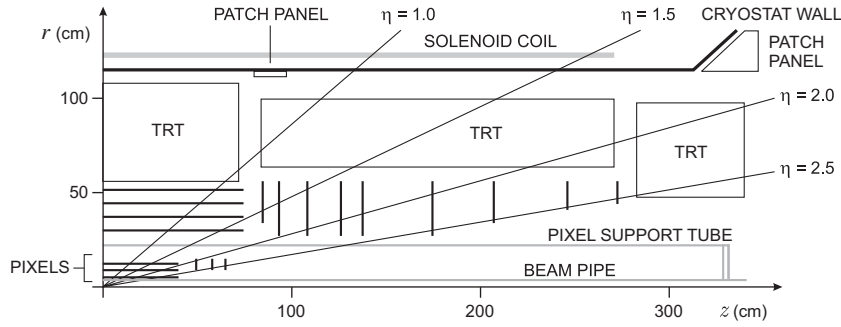


**Fig. 7.** Hadron (top) and neutron (bottom) fluences over one year in the ATLAS inner detector. The vertical scale ranges from  $10^{12}$  to  $10^{14}\ \text{cm}^{-2}$ . Adapted from Ref. [15].

power to the individual detector modules and the required cooling. The material in the power cabling can be reduced to some degree, but if its resistance is too high, the power dissipated in the cabling can be significant and furthermore failure of a module within a group can increase the voltage to an unacceptable level that makes the other modules fail.

Reducing the required power in the local readout electronics is a major challenge. Since many portions of the detector are not readily accessible and actually require major disassembly, reliability is a major issue. Unlike many previous detectors where somewhat immature designs were incorporated with the assumption that flaws would be found and then repaired, the large-scale subsystems in LHC detectors had to be thoroughly tested before proceeding into mass production. Furthermore, since production had to be distributed over multiple institutions, often in various continents, specifications and production techniques had to be well established to ensure uniform results. Frequently, testing procedures focus only on expected problems, but it is even more important to consider unexpected flaws and “reliable” components should not be ignored.

Finally, these detectors turned out to be some of the most expensive ones ever made. Although some cost estimates only include material, in reality one has to include effort. This is not just a cost driver that should promote the development of efficient designs, but a major constraint is that these systems require considerable experience and understanding, and the number of people up to this is limited. In this respect large



**Fig. 8.** Cross-section through one quadrant of the ATLAS SCT showing the arrangement of the tracking detector subsystems. The silicon tracker is about 1 m in diameter and 5.4 m long. The pixel subsystem is self-contained and can be inserted or removed separately. Intermediate patch panels mounted at the cryostat wall facilitate assembly. Detailed dimensions are given in Ref. [17].

collaborations are an advantage, as one can select competent contributors from different institutions to form individual working groups. Since many participants are driven by specific experiences and stereotypes, going beyond the perspective of a single institution can be essential.

## 5. LHC Detector configurations and subdetector signal sources

As noted above, not all interesting aspects of the different LHC detectors can be presented. The following section provides an overview of the types of subdetectors used in ATLAS, CMS, LHCb, and ALICE to illustrate the various signal sources and some different aspects of signal processing.

### 5.1. ATLAS

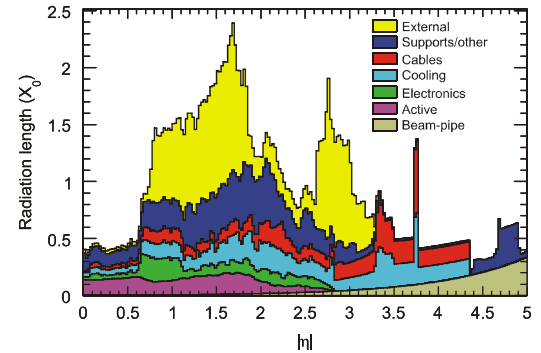
ATLAS [16] consists of a tracking system, enclosed in a 2 T solenoid, electromagnetic and hadron calorimeters, and an extensive muon detector with high position resolution.

The ATLAS tracking system consists of three parts, the silicon pixel detector (5–12 cm radius), the semiconductor tracker (SCT, 30–51 cm radius) consisting of silicon strip detectors, and the transition radiation tracker (TRT, 56–107 cm radius) as shown in Fig. 8. In addition to the main barrel component, all three have endcap components (e.g. disks in the semiconductor detectors) that together extend tracking to a rapidity  $\eta = 2.5$ . The semiconductor detectors will be discussed in more detail in Sections 8 and 9. The TRT utilizes 4 mm gaseous straw tubes of 144 cm length, arranged parallel to the beam axis in the central region and arranged radially with 37 cm length in the endcap regions. The number of readout channels in the pixel detector is 80.4 million, in the SCT 6.3 million, and in the TRT about 350 000. Fig. 9 shows the material distribution of the various infrastructure components.

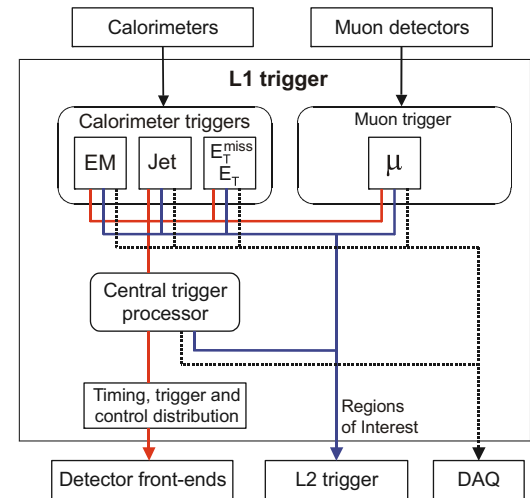
The electromagnetic calorimeter is a liquid argon (LAR) ionization chamber utilizing Pb absorbers. The electrodes are accordion shaped to provide  $\Phi$  symmetry without azimuthal cracks. Charge velocity in liquid argon is low, so the required short signal pulses are obtained by a combination of electrode geometry and special pulse shaping, while still maintaining the signal resolution. This is discussed in Section 6.4.

The hadron calorimeter is placed directly outside the EM calorimeter. It is a sampling calorimeter using steel absorber layers and scintillating polystyrene tiles as active material. The scintillator tiles are read out by wavelength-shifting optical fibers, which are coupled to photomultiplier tubes. The hadronic endcap and forward calorimeters utilize liquid argon.

The muon system is often shown in photographs because of the novel toroids that shape the magnetic field to optimize the



**Fig. 9.** Material distribution of the ATLAS tracker shown for various infrastructure components, i.e. support, cooling, cables [16]. The active sensor material and electronics only contribute a fraction of the total material.



**Fig. 10.** Components of the ATLAS L1 trigger combining data from the calorimeters and muon detectors [16].

position resolution. Over most of the rapidity range the detectors are rather conventional drift tubes. At large pseudorapidities, multiwire proportional chambers with cathodes segmented into strips provide the desired position resolution. Trigger information is gained from an additional set of resistive plate chambers covering  $\eta < 2.4$ , which provide bunch-crossing identification,  $p_T$  thresholds, and the muon coordinate orthogonal to that provided by the precision-tracking chambers.

The Level 1 (L1) trigger is derived from the calorimeters and muon detectors. This initiates readout of the tracking data, which

is then used for higher trigger levels. Fig. 10 shows the Level 1 block diagram and Fig. 11 shows the processing blocks used to derive the jet and cluster information used to derive the L1 trigger. Fig. 12 shows an example of how the information from the ATLAS muon detectors is processed to contribute to the L1 trigger.

Overall, all of these detectors required careful readout designs, but the most demanding designs are in the silicon pixel and

strip detectors, and the ATLAS liquid Ar electromagnetic calorimeter.

## 5.2. CMS

CMS [18] is also a general purpose detector. The CMS tracker is all-silicon, consisting of a pixel system and a large strip detector system covering about 210 m<sup>2</sup>. As shown in Fig. 13 the barrel region includes three layers of pixel detectors and 10 layers of silicon strip detectors. The CMS tracking system is enclosed in a 4 T solenoid. Fig. 14 shows the material distribution of the various infrastructure components. A more detailed description will follow in Section 8.

The EM calorimeter uses PbWO<sub>4</sub> scintillators read out by silicon avalanche photodiodes (APDs) in the barrel region and vacuum phototriodes (VPTs) in the endcaps. The hadron calorimeter utilizes brass/scintillator sampling, where the scintillation light is extracted from the scintillator tiles by wavelength-shifting fibers, which are coupled to clear fibers that couple to hybrid photodiodes, which provide gain and can also operate in high axial magnetic fields. For rapidities  $3 < \eta < 5$  a calorimeter using iron/quartz fiber technology is used. The quartz fibers emit Cherenkov light that is measured by photomultipliers.

The muon detectors utilize three different types of gaseous detectors. In the barrel region the muon rate is low and drift chambers with rectangular drift tubes are used. Rates in the endcaps are much higher and the magnetic field is large and non-uniform, so cathode strip chambers with fast response times are used. To deal with potentially increased background rates and improve the accuracy in recording the beam crossing time, resistive plate chambers are incorporated in both the barrel and endcap regions. Operating in avalanche mode they have fast response with good time resolution, but the barrel drift chambers and endcap cathode strip chambers provide the required position resolution.

The barrel chambers use custom integrated circuits with charge-sensitive amplifiers providing a 100  $\Omega$  input impedance in the range 5–200 MHz. Input protection has been implemented to soak up the charge that can be released by multiple sparks from the 470 pF capacitors that couple the 3.6 kV biased wires to the preamp input. Fig. 15 shows the block diagram of the four-

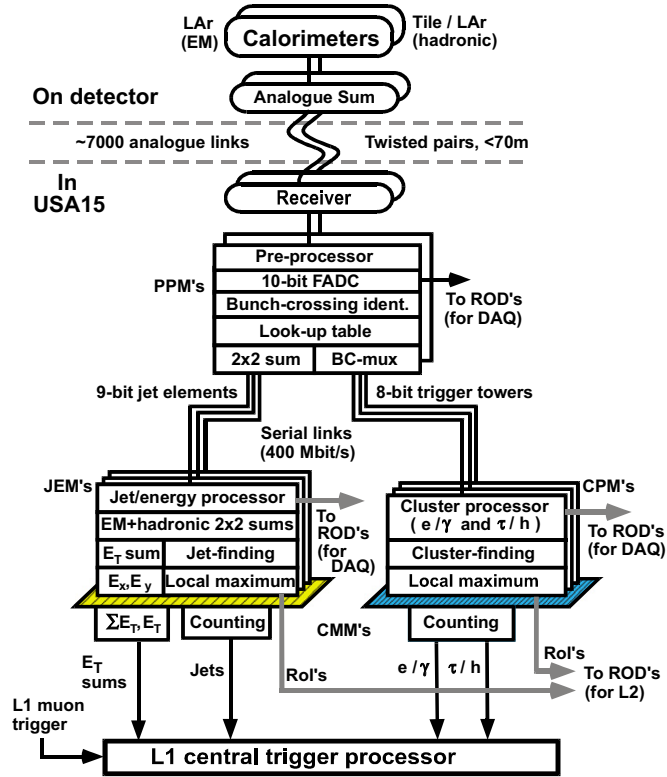


Fig. 11. Processing sequence of the ATLAS L1 calorimeter trigger system [16].

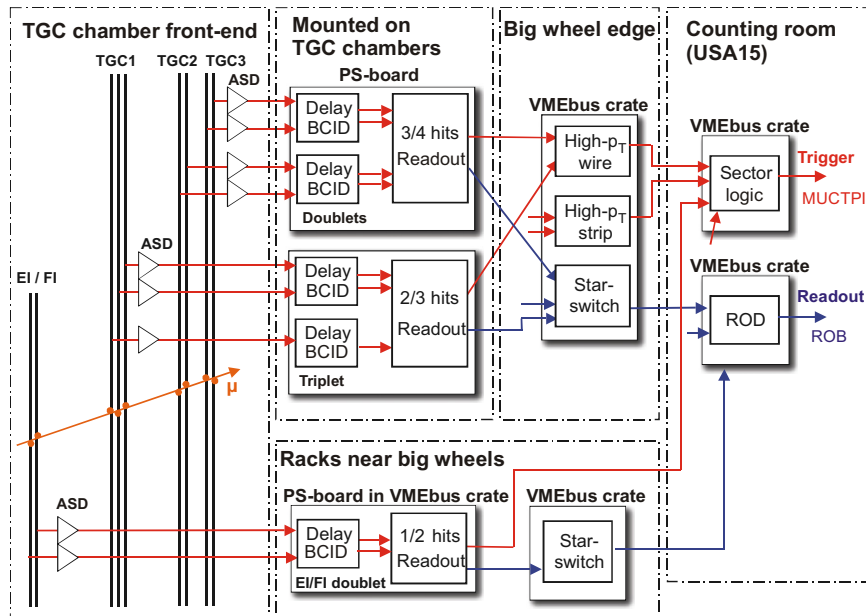


Fig. 12. Block diagram of the muon section of the ATLAS L1 trigger [16].



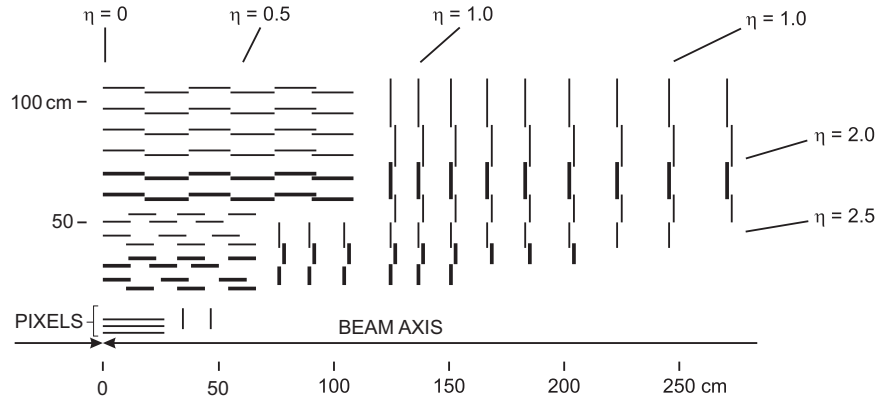


Fig. 13. Cross-section through one quadrant of the CMS tracker showing the arrangement of the strip and pixel systems.

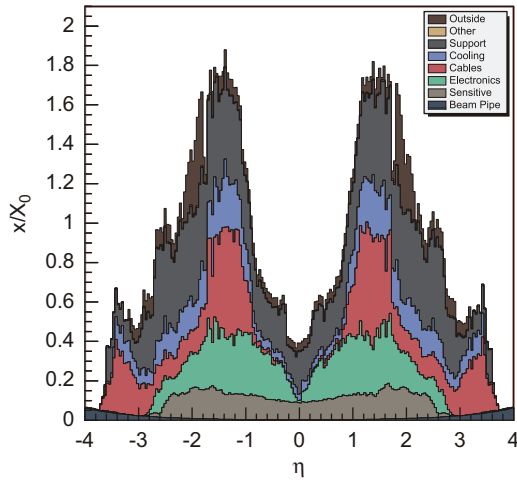


Fig. 14. Material distribution of the CMS tracker shown for various infrastructure components, i.e. support, cooling, cables [18]. The active sensor material and electronics only contribute a fraction of the total material.

channel readout chip. Balanced LVDS output drivers are crucial in suppressing digital cross-talk to the analog inputs.

The endcap cathode strip chambers use both anode and cathode readouts. The anode custom IC has 16 channels of preamplifiers and 30 ns semi-Gaussian shapers with tail cancellation to suppress the slow signal component due to positive ions drifting away from the anode wires [19]. With 30 ns shaping the observed signal contains about 12% of the total avalanche charge, yielding an average of 130 fC, which together with the 1.4 fC noise level yields a high signal-to-noise ratio. A constant-fraction discriminator provides timing with less than 3 ns time walk. The hit signals are fed to a digital processing board that uses an FPGA to search for hit patterns among the multiple planes that are consistent with tracks originating from the interaction region.

The cathode readout operates with a 100 ns shaping time, also with tail cancellation, and stores the analog signals in a switched capacitor array that feeds a 12-bit ADC. The analog signal also feeds a comparator that together with its neighbors can promptly identify a muon hit location to within one-half of a strip width.

### 5.3. LHCb

LHCb [27] is dedicated to heavy flavor physics, aimed toward indirect evidence of new physics in CP violation and rare decays. Tracking is a key component together with short latency triggering capable of efficient event selection. In order to measure track coordinates close to the interaction the vertex detectors are

arranged as circular disks that are split into two halves so that the detectors can be retracted during beam injection. A subset of the vertex detector layers is shown in Fig. 16.

The front-end architecture is designed for a first level trigger rate up to 1 MHz, an order of magnitude higher than ATLAS or CMS. The front-end configuration is shown in Fig. 17. Detector signals are sampled at a 40 MHz rate and stored in 4  $\mu$ s deep pipeline memories while the first level trigger selection is made. For a more detailed description and further references see Ref. [27]. Radiation damage levels are similar to those in the pixel detectors of ATLAS and CMS.

The VELO front-end IC (“Beetle” chip) fabricated in 0.25  $\mu$ m CMOS [28] consists of a charge-sensitive amplifier with subsequent gain stages that yield an output pulse with about 20 ns peaking time. Note that since signals are synchronized with the beam crossing frequency, sampling these signals at a 40 MHz rate yields the correct amplitudes. Random pulses would require a much faster sampling rate. A 160-stage analog pipeline provides the 4  $\mu$ s latency time. Analog signals are sent to the external readout system where the signals are digitized and preprocessed, i.e. applying zero-suppression, pedestal subtraction, cross-talk removal, channel reordering, common mode suppression, and reordering. Zero-suppressed data are sent to the trigger system.

Downstream from the vertex detector, silicon strip detectors provide tracking information. The detector layers are subdivided into readout sectors depending on the event rate, but also aimed at minimizing the number of readout channels, as this is a major cost component.

The outer tracker utilizes gaseous straw-tube modules. Each drift-tube has an inner diameter of 4.9 mm. Custom eight-channel front-end chips include the complete analog chain with a threshold discriminator. Utilizing the DMILL process, the bipolar transistor front-end yields a peaking time of 7–8 ns. An additional IC utilizing the 0.25  $\mu$ m CMOS process feeds the signals from four front-end ICs to 32 time digitizers, which use 64 time-delay stages to yield about 400 ps time resolution.

Particle identification is derived from RICH detectors. The upstream detector RICH1 covers the low-momentum particle range 1–60 GeV/c using aerogel and C<sub>4</sub>F<sub>10</sub> radiators. The downstream detector covers higher momenta from 15 to > 100 GeV/c using a CF<sub>4</sub> radiator. Hybrid photon detectors are used to detect the Cherenkov photons in the wavelength range 200–600 nm. These detectors provide internal gain by accelerating the photoelectrons emitted from the photocathode to energies of 10–20 keV and detecting them in a silicon 1024-pixel hybrid array, which provides both position resolution and reduced noise levels that improve optical sensitivity. This is a concept that has existed for some time, but actually bringing it to practical use is

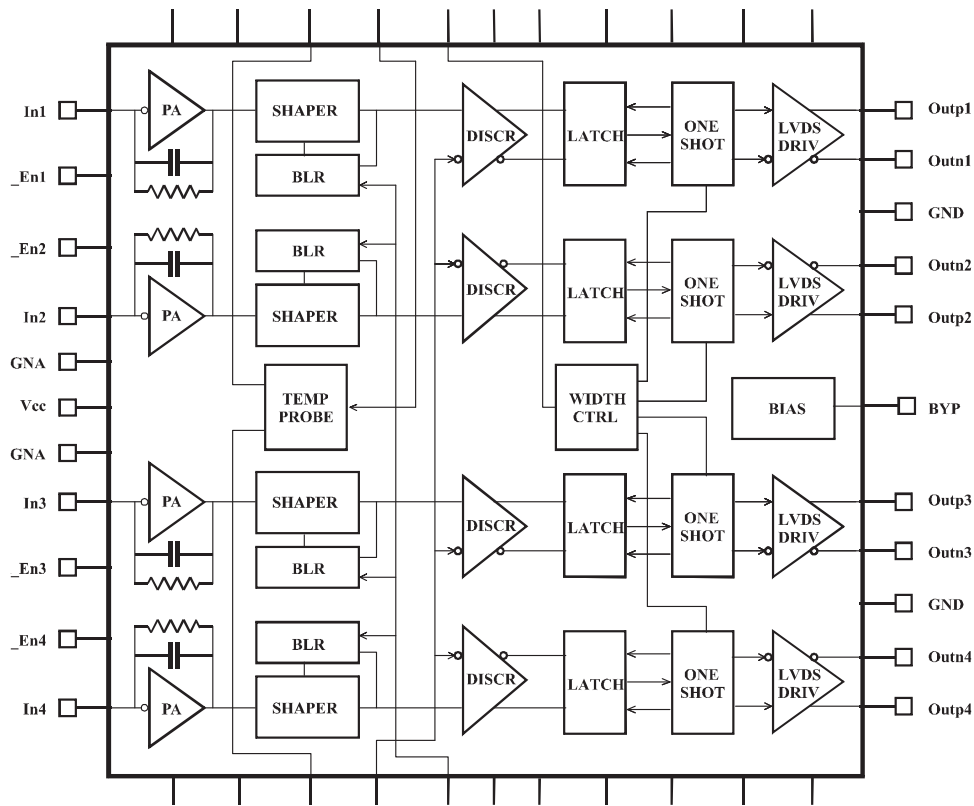


Fig. 15. Block diagram of the CMS drift-tube muon chamber readout IC [18].



Fig. 16. Half-disks with front-end electronics used in the LHCb VELO vertex detector [29].

another important accomplishment brought about by large detector projects.

The electromagnetic and hadron calorimeters both use scintillators coupled through wavelength-shifting fibers to phototubes. The signals are shaped prior to digitization sampled at 40 MHz with 12-bit flash ADCs. Fig. 18 shows the traditional form of PMT pulse shaping using delay-line clipping both at the PMT anode and the preamplifier input. The input buffer amplifier and charge integrator are implemented in 0.8  $\mu\text{m}$  BiCMOS.

The muon system consists of five rectangular units along the beam axis covering a total area of 435  $\text{m}^2$ . The detectors are multi-wire proportional chambers with electronics designed to maintain 5 ns time resolution. Custom designed ICs are used for the readout, enclosed in standard packages that allow conventional PC boards that are plugged directly onto the chambers (Fig. 19). The front-end chips each have eight channels

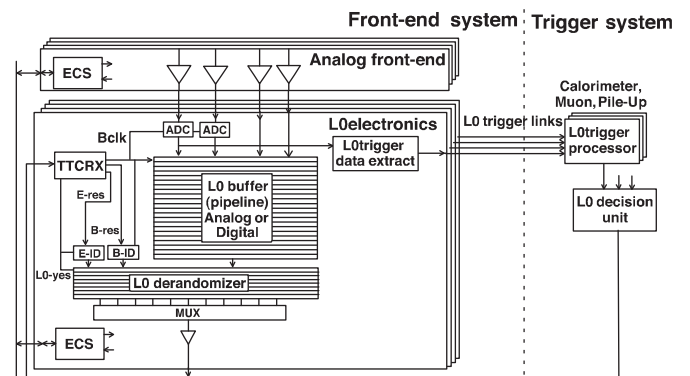


Fig. 17. LHCb front-end architecture to enable fast Level 1 triggering. Adapted from Ref. [27].

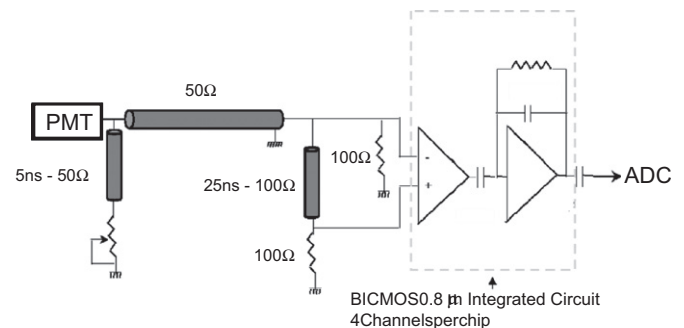
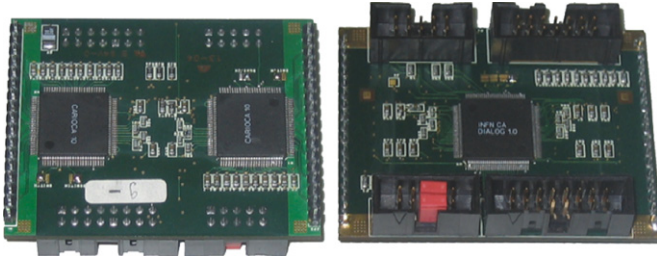


Fig. 18. Front-end of the LHCb calorimeter readout [27].

of front-end current amplifiers with 50  $\Omega$  input impedance, and discriminators with individually adjustable thresholds. Two front-end chips feed a data pre-processing chip that performs



**Fig. 19.** Top and bottom views of the LHCb muon readout boards [27]. The two front-end chips are shown on the left and the data pre-processing chip on the bottom side of the right.

the logical OR of corresponding pads in the two layers of a chamber and also includes individual 1.6 ns delay adjustments to match timing between channels. Individual channels can also be masked.

#### 5.4. ALICE

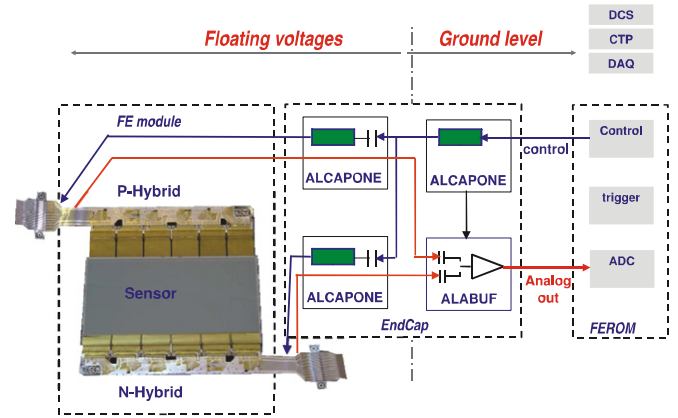
ALICE [30] is a heavy-ion detector. It will run at lower luminosity and have a lower event rate than the other detectors ( $10^4 \text{ s}^{-1}$  for Pb–Pb collisions), but with a much higher multiplicity per collision. ALICE design targeted  $dN/d\eta = 4000$  while evaluating performance at twice the multiplicity. The radiation dose is moderate ( $< 3 \text{ kGy}$ ) and together with the lower event rate it allows slow and rather conventional detectors. However, segmentation must be high to distinguish the large number of simultaneous tracks from one another. Furthermore, particle identification is essential, so multiple techniques are used:  $dE/dx$ , time of flight (TOF), transition and Cherenkov radiation, EM calorimetry, muon filters, and topological decay reconstruction. For a detailed description and further references see Ref. [30].

The central portion of the detector is enclosed in a solenoid magnet, which includes six planes of Si pixel (SPD), drift (SDD), and strip (SSD) detectors, a cylindrical time projection chamber (TPC), three particle identification arrays using time-of-flight (TOF), ring imaging Cherenkov detectors (HMPID), transition radiation detectors (TRD), and two electromagnetic calorimeters (PHOS and EMCal). All detectors except HMPID, PHOS, and EMCal cover the full azimuth. In addition, a forward arm covering angles 2–9 degrees includes absorbers, a large dipole magnet, and 14 planes of tracking and triggering chambers. A scintillator array at the top of the magnet is used to trigger on cosmic rays.

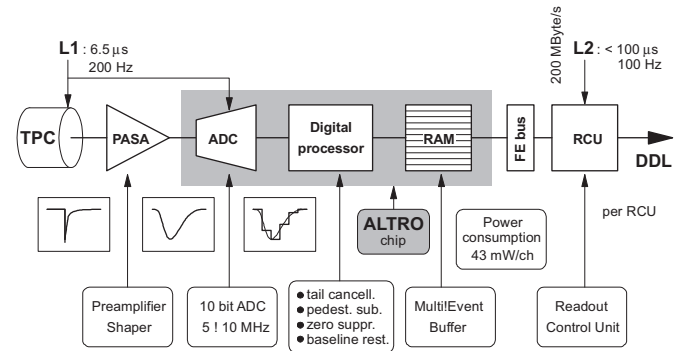
Si pixel detectors form the innermost two tracking layers (3.9 and 7.6 cm radius) with about 9.8 million channels. The following two layers use silicon drift detectors (15.0 and 13.9 cm radius) with about 133 000 channels, and the outer two silicon layers (38 and 43 cm radius) are double-sided strip detectors with 2.6 million channels. Maximum occupancies are 2.1% in the pixel layers, 2.5% in the silicon drift layers, and 4% in the silicon strip layers. The maximum radiation dose is 270 krad (2.7 kGy), imposed on the inner pixel layers.

The pixel detectors are bump-bonded hybrids with a binary readout. The readout ICs are fabricated using the IBM 0.25  $\mu\text{m}$  CMOS process [20,21], as used for the ATLAS pixels and the CMS pixel and strip systems. Data are transmitted to the off-detector readout electronics by optical fibers.

The silicon drift detector front-end chips combine preamplifiers, pulse shapers with a 40 ns peaking time, analog storage pipelines, and a set of 10-bit successive approximation ADCs that digitize the data in time windows identified by the trigger [22–24]. Additional chips perform digital processing and data compression before data are sent from the detector.



**Fig. 20.** Arrangement of the ALICE Si-strip readout [30]. The detector modules combine sensors and front-end chips in the active volume. The endcap is located at the end of the detector ladders and the FEROM is located outside the magnet.



**Fig. 21.** Block diagram of the ALICE TPC readout [30].

Fig. 20 shows the readout arrangement for the Si strip detectors. The front-end chip mounted on the detector module includes 128 channels consisting of preamplifiers that can accept both positive and negative input charge, pulse shapers with adjustable 1.4–2.4  $\mu\text{s}$  shaping times, and peak detectors that sample and hold the peak amplitude, which is transferred through a multiplexer to the output line. The endcap chips perform control functions and also include analog drivers to send the signal to the FEROM outside the magnet, where the signals are digitized and then sent to the off-detector electronics.

The TPC is the major tracking device, providing three-dimensional track reconstruction and  $dE/dx$  resolution better than 5–7%. Its endplates are rather conventional multi-wire proportional chambers with cathode pad readout. The wire geometry and pad size are modified between the inner and outer radii to adapt to the track density. The signals are read out by custom ICs that for each channel include a charge-sensitive amplifier and preshaper. The output feeds a second chip with 10 bit 25 MHz ADCs running at a lower sampling rate as indicated in Fig. 21, which reduces power consumption. Integrated with the ADCs are digital pulse processors that perform tail cancellation, pedestal subtraction, zero-suppression, and baseline restoration. Fig. 21 shows the block diagram of the complete readout.

The transition radiation detector is sensitive to electrons with momenta  $> 1 \text{ GeV}/c$ . Together with the specific energy loss in an appropriate gas mixture the transition radiation can be used to reject pions and provide electron identification. It is part of the Level 1 trigger. Subdivided into 6  $\text{cm}^2$  pads, the maximum occupancy is 34% for the simulated multiplicity density of  $dN/d\eta = 8000$ . The detector elements consist of a sandwich

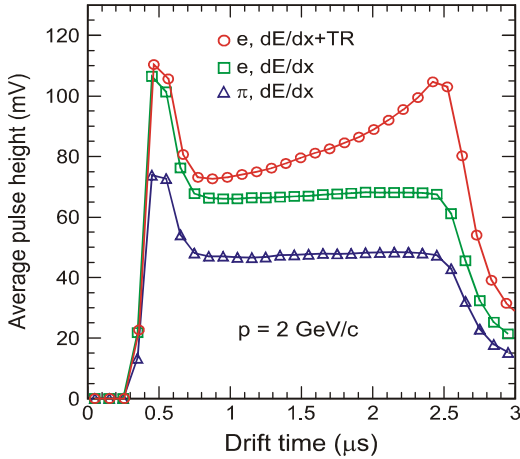


Fig. 22. TRD pulse shapes for electrons and pions [30].

radiator of 48 mm thickness, a 30 mm drift section, and a multi-wire proportional chamber with pad readout. Ionizing radiation releases electrons in the counting gas, but particles exceeding the transition radiation threshold produce X-ray photons with energies of 1–30 keV. Electrons can be distinguished from pions by analyzing the pulse shape as shown in Fig. 22. A charge-sensitive amplifier plus pre-shaper with 120 ns shaping time feeds the signal into a 10 bit ADC with a 10 MHz sampling rate. Digital signal processing separates the different pulse shapes.

The time-of-flight is measured with a large-coverage detector with more than  $10^5$  independent channels to keep the occupancy below 15%. This is achieved with a multi-gap resistive plate chamber. This is not a drift chamber, so the time jitter is dominated by fluctuations in the formation of the avalanche and the pulse shape is better suited for timing. Initial tests were done with commercial readout chips, but the final design uses custom 0.25  $\mu\text{m}$  CMOS ICs with substantially lower power dissipation (40 mW per channel).

Photons spanning the range from thermal to QCD processes and neutral mesons are measured in a rather small, but highly segmented PbWO<sub>4</sub> scintillation calorimeter. A set of multiwire chambers in front of the EM calorimeter are used to veto charged particles. An additional EM calorimeter (EMCal), a Pb-scintillator sampling system, will measure jet rates and fragmentation functions in connection with tracking data.

A forward muon spectrometer is designed to measure the production of heavy-quark resonances with sufficient mass resolution to separate all states. It consists of a composite absorber ( $\approx 10\lambda$ ) made of layers of high- and low-Z materials within a 3 T magnet. Ten planes of thin and high-granularity cathode-strip tracking stations are used to reconstruct the tracks. A second muon filter consisting of about  $7\lambda$  of iron with a subsequent four planes of resistive plate chambers is used for muon identification and triggering.

Two sets of 12 Cherenkov detectors utilizing fine mesh photo-multipliers with fused silicon radiators measure the event time to a precision of 25 ps. A number of small and specialized detectors are used for triggering. Each pixel chip provides a digital pulse when at least one pixel is hit. These outputs are transmitted every 100 ns on 120 optical links. Two-way optical splitters allow the simultaneous transmission of trigger and readout data in the optical fiber links. Two arrays of segmented scintillators are used for the minimum bias trigger. Forward charged particle multiplicity detectors at  $-3.4 < \eta < -1.7$  and  $1.7 < \eta < 5$  using silicon strip detectors and photon multiplicity detectors at  $2.3 < \eta < 3.7$  using scintillators with quartz fiber outputs provide additional trigger information. Two sets of compact calorimeters provide the impact parameter

trigger using scintillators with embedded quartz fibers for the signal transfer. The scintillator array at the top of the detector is used for calibration and alignment tests using cosmic rays.

The ALICE trigger operates at several levels to utilize the timing properties of various detectors. A pretrigger activates the TRD at about 900 ns after each interaction. Subsequently, the L0 and L1 reduce the event rate after 1.2 and 6.5  $\mu\text{s}$ , respectively. The final L2 trigger is issued after about 100  $\mu\text{s}$ , to incorporate the maximum drift time of the TPC.

## 6. Front-end configurations

### 6.1. Performance specifications

#### 6.1.1. Electronic noise

The electronic noise charge has two contributions, the first due to current noise, which increases with integration time, i.e. the pulse area, and the second due to voltage noise, which translates to charge through the total passive input capacitance  $C$  and increases with the bandwidth of the shaper, so it increases with decreasing shaping time, i.e. higher frequency Fourier components. The equivalent noise charge

$$Q_n^2 = i_n^2 F_i T_S + e_n^2 F_v \frac{C^2}{T_S} + F_{vf} A_f C^2 \quad (4)$$

where  $T_S$  is a characteristic time parameter, e.g. the peaking time of a pseudo-Gaussian pulse. The parameters  $F_i$  and  $F_v$  are determined by the pulse shape:

$$F_i = \frac{1}{2T_S} \int_{-\infty}^{\infty} [W(t)]^2 dt \quad (5)$$

and

$$F_v = \frac{T_S}{2} \int_{-\infty}^{\infty} \left[ \frac{dW(t)}{dt} \right]^2 dt. \quad (6)$$

Since both parameters are normalized to  $T_S$ , its definition is somewhat arbitrary, but it must perform the proper scaling for both current and voltage noise, so it must characterize both the pulse area and the derivatives, i.e. the rise and fall times.

The third term is the contribution of “ $1/f$ ” noise. Typically an equivalent voltage source, its contribution increases with the total capacitance at the input, but is independent of  $T_S$ , as the total contribution of a  $1/f$  noise spectrum is determined by the ratio of the upper to lower cutoff frequency, rather than the bandwidth. For a given shaper configuration this ratio is independent of shaping time. The “ $1/f$ ” contribution is typically significant at the optimum shaping time, where the current and voltage contributions are equal. However, when fast timing and increased shot noise due to radiation damage are important, the selected shaping time is usually smaller, so the voltage noise dominates. Then the “ $1/f$ ” noise contribution is usually negligible, as the components add in quadrature. It should also be noted that low-frequency noise often does not have a  $1/f$  frequency dependence, but can have multiple components with a frequency dependence  $dP_n/df = 1/f^\alpha$ , where  $\alpha$  is in the range 0.5–2 (see Ref. [1]).

The first two noise contributions are often called “parallel” and “series” noise, but viewing them in terms of physics quantities as noise current and voltage indicates their behavior. Current noise is due to statistical fluctuations in current flow, so the absolute fluctuations in charge increase with integration time, whereas voltage noise behaves like thermal noise, so its contribution increases with bandwidth and translates to charge through the total capacitance at the input. For a more detailed discussion see Ref. [1].



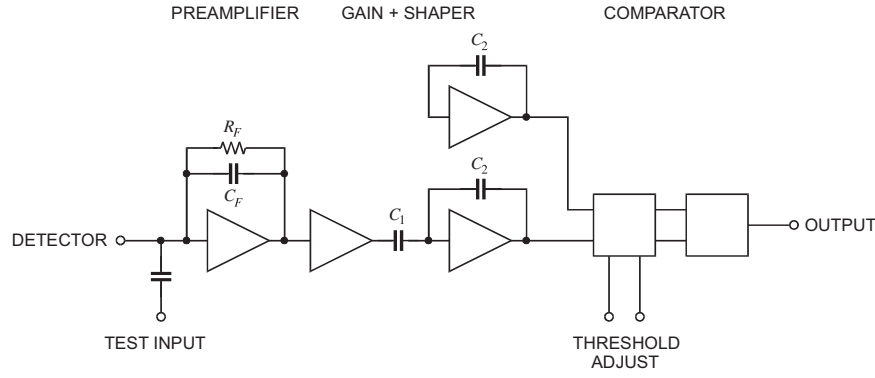


Fig. 23. Block diagram of a charge-sensitive amplifier feeding gain/shaping stages and a threshold discriminator.

Minimizing noise under conflicting conditions involves a careful choice of shaping times. Radiation damage increases the detector leakage current, so reducing the shaping time will reduce its contribution. For a given peaking time, a symmetrical pulse yields a smaller current noise contribution, as the area will be smaller for a given slope, so multiple integrators are desirable. However, as will be shown below, this does not have to increase circuit complexity. Although short shaping times reduce the sensitivity to current noise, they increase the voltage noise contribution, as this increases the frequency range. The noise current typically originates from the detector leakage current, the base current of a bipolar input transistor, and the resistor shunting the feedback capacitor in the charge-sensitive amplifier. Both the detector leakage current and the base current are sensitive to radiation damage. The origin of the voltage noise is typically the preamplifier's input transistor, so if its operating point changes with radiation damage, the overall noise will also be affected.

Although short shaping times reduce sensitivity to current noise and improve time resolution, which can be quite desirable, if the peaking time is less than the width of the signal current pulse from the detector, part of the signal will be lost ("ballistic deficit"). If the overall peaking time of the electronics is comparable to the width of the signal pulse from the detector, a change in peaking time will change the charge response. Alternatively, if the detector pulse width is reduced by increasing the detector bias, the measured signal can increase, although the same charge was deposited in the detector.

Since the relative noise contributions depend on the shaper parameters  $F_i$ ,  $F_v$ , and  $T_s$  (when " $1/f$ " noise is negligible), all three must be specified to characterize a pulse shaper. In circuits designed to reduce power dissipation, the shaping times typically depend in part on the bandwidth of the preamplifier, which in turn depends on the load presented by the detector. Thus, simply specifying the noise level alone for a given configuration will not necessarily predict the results for another configuration. Furthermore, to assess the effects of radiation damage, the individual noise contributions must be analyzed. Since changes in detector parameters or electronic operating points can affect all three shaper parameters  $F_i$ ,  $F_v$ , and  $T_s$ , they should be characterized for various detector capacitances and electronics operating points. It is also important to verify amplifier stability, as "ringing" in the preamplifier can affect the signal's peak amplitude at the shaper output.

### 6.1.2. Input impedance

Controlling the input impedance is critical when an amplifier receives signals through a long cable and reflections are to be

minimized. However, in strip and pixel detectors the preamplifier input impedance is also important, as it sets the cross-talk between adjacent electrodes. To capture the bulk of the signal charge in the target channel, the amplifier's input impedance must be small compared to the impedance presented by the inter-electrode capacitance  $C_{ss}$ . This impedance  $1/\omega C_{ss}$  depends on the range of frequencies at which the signal amplitude is measured, i.e. the shaping time.

At low frequencies, where the amplifier's open-loop phase shift is  $180^\circ$ , the input impedance of a charge-sensitive amplifier utilizing capacitive shunt feedback is capacitive. However, the open-loop corner frequency where the gain begins to roll off and introduces a  $90^\circ$  phase shift is typically well below the passband set by the shaper, approximately centered at the inverse peaking time. There the input impedance levels off and is resistive (phase shift  $0^\circ$ ). See Fig. 23 for the amplifier configuration and Ref. [26] for a more detailed discussion. The value of the resistive input impedance

$$Z_i = \frac{1}{\omega_0 C_f} \quad (7)$$

depends on the feedback capacitance  $C_f$  and the amplifier's unity gain frequency  $\omega_0$ , extrapolated from the frequency regime where the gain falls off linearly with frequency.

Cross-coupling between strips also introduces additional noise from the neighbors (see Ref. [1] for a more detailed discussion). Although purely capacitive cross-coupling between strips is often assumed, after radiation damage the interstrip impedance can also have a resistive component. Charge buildup at the surface between strips can introduce states with a lifetime and this will introduce a lossy (ohmic) component at frequencies depending on the lifetime. Again, the key frequencies are determined by the peaking time, so the interstrip impedance should be measured as a function of frequency to verify that capacitive coupling dominates.

Although some parameters can be derived *in situ* and some papers show the pulse shape, which can be used to estimate  $F_i$  and  $F_v$ , it would be better to specify these parameters directly, since they are easy to extract from the circuit simulations. A plot of the input impedance magnitude and phase vs. frequency should be included, showing both the simulation and measurement results. The relevant frequency range can be derived from the pulse shape, but showing the front-end's overall frequency response would also be useful.

In strip and pixel detectors the input impedance at the peaking frequency is critical. The low-frequency gain (kHz range) determines baseline stability, but not the input impedance in the shaper passband. The magnitude and phase of the input impedance should be measured as a function of frequency.



A consistency check is provided by the measured open-loop gain and phase of the preamplifier vs. frequency. The measured data should be compared with simulations.

## 6.2. Charge-sensitive front-ends and shapers

Fig. 23 illustrates a charge-sensitive front-end. This example shows a binary readout, but the considerations in optimizing the preceding gain stages also apply to analog readouts. In the “standard” configuration the charge-sensitive preamp and the shaper are viewed as separate functions, i.e. the preamp output has a risetime that is much smaller than the peaking time at the shaper output and the decay time of the preamp output signal is much longer than the shaped pulse. However, for fast peaking times and where power consumption must be minimized, the preamp and shaper functions can be combined. As noted above, reducing the rise time increases the voltage noise contribution, so the rise time is chosen to just meet the timing requirements. The current noise contribution increases with the area of the shaped pulse, so increasing the symmetry of the shaped output is desirable. This is achieved by multiple integrators.

In the overall pulse shaping the decay time is set by the discharge time constant of the preamplifier feedback capacitor  $\tau_d = R_F C_F$ . The rise time at the shaper output results from the bandwidths of the preamplifier and the two subsequent gain stages. The gain stages are necessary to provide sufficient gain so that the threshold voltage at the comparator is sufficiently high to provide good channel-to-channel and chip-to-chip uniformity. Since the gain stages are necessary, utilizing them as pulse shapers does not increase the power requirements.

Since the first two stages are single-ended to reduce power consumption, substantial circuit complexity would be necessary to maintain DC stability, so AC coupling is introduced at the input of the third stage. From here on the circuitry is differential. The third stage is still single-ended, but it is replicated as a dummy amplifier to bias the second input of the differential comparator. The dummy amplifier is to be included in each individual channel to obtain optimal parameter tracking under radiation damage, and also to maintain a parallel architecture and reduce single-point failure modes. The threshold level is applied differentially to exploit device tracking during irradiation. In modern designs this is digitally controlled by a digital-to-analog converter (DAC). The calibration circuitry (test input) provides a means to monitor the gain. Similar principles have been applied to bipolar transistor and CMOS circuitry. Maintaining threshold matching under irradiation is a challenge, so it is sometimes necessary employ trim DACs on each channel to correct the master threshold level channel-by-channel, as shown below in the ATLAS silicon pixel and strip readouts.

In analyzing the effect of changing detector parameters on the front-end performance, differences in capacitance do not only affect the translation of noise voltage into noise charge. In a low-power design they can also affect the shaping time parameters, which can further contribute changes in noise and also affect the sensitivity to radiation damage. That is why specifying the shaper parameters is necessary to identify which system parameters contribute most to the effects of radiation damage.

## 6.3. Low-power front-ends

The second term in Eq. (4) is the noise voltage contribution, which is inversely proportional to shaping time, but also depends on a detector parameter, namely the capacitance. The spectral noise voltage density  $e_n$  is typically dominated by the noise of the amplifier input transistor. Reducing the spectral noise voltage density of the transistor requires an increase in current. Under

optimum scaling the power in the input stage increases with  $(S/N)^2$ , e.g. if the signal is reduced to 1/2, maintaining the signal-to-noise ratio requires at least a four-fold increase in power. For a more detailed discussion see Ref. [1].

Assume that the voltage noise dominates and radiation damage has reduced the signal level to  $S_{\text{rad}}/S_0$ . Then reducing two parameters can reduce the noise, i.e. the amplifier's spectral noise density  $e_n$  or the detector capacitance  $C$ . Reducing  $e_n$  requires that for a given detector capacitance the power in the input transistor would have to be increased by  $(S_0/S_{\text{rad}})^2$ . On the other hand, reducing the capacitance, e.g. the length of strip electrodes, can reduce the noise proportionally. However, reducing the strip length by  $S_{\text{rad}}/S_0$  increases the number of readout channels by  $S_0/S_{\text{rad}}$ , so to maintain  $S/N$  for 1/2 of the signal level, the input stage contributions increase 2-fold, rather than 4-fold if the strip length is kept constant and the input stage power is increased. Generally, increasing segmentation is the efficient path toward reducing power with increased radiation damage, as reducing the pixel or strip area also reduces the detector's current noise contribution.

However, front-end power is not determined by the analog circuitry alone. The digital circuit blocks can also contribute significantly. Digital power scales with the logic levels, i.e. the digital supply voltage  $V_{DD}$  and the clock frequency  $f$ :

$$P_{\text{dig}} \propto V_{DD}^2 \cdot f \quad (8)$$

so the reduced supply voltage associated with smaller CMOS feature sizes is advantageous.<sup>1</sup> However, increasing the readout speed, i.e. the clock frequency, increases the power proportionally.

Reducing strip length  $L$  reduces the power in the input stage, but increasing the number of chips increases the digital power contribution. The total power for  $N_{\text{strip}}$  readout channels with an analog power of  $P'_{\text{analog}}$  per unit length and a digital power per channel  $P_{\text{dig}}$  is

$$P_{\text{tot}} = N_{\text{strip}}(P'_{\text{analog}}L^2 + P_{\text{dig}}). \quad (9)$$

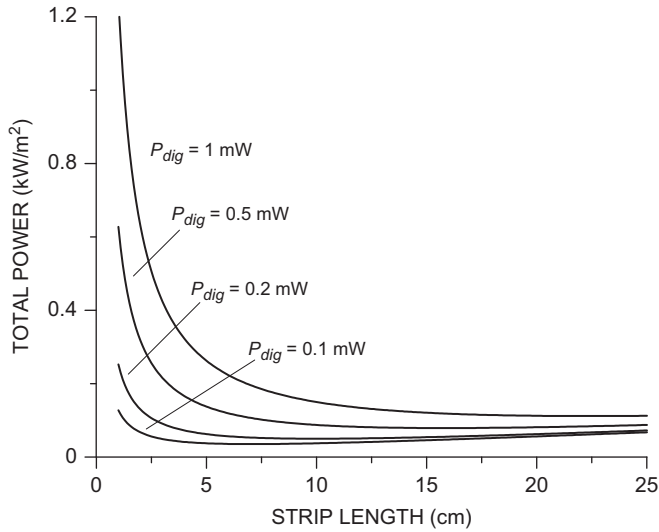
The number of strips for a total detector area  $A$  with a strip length  $L$  and pitch  $p$  is  $N_{\text{strip}} = A/pL$ , so the total power per unit area

$$\frac{P_{\text{tot}}}{A} = \frac{1}{p} \left( P'_{\text{analog}}L + \frac{P_{\text{digital}}}{L} \right). \quad (10)$$

Fig. 24 shows the total power per  $m^2$  vs. strip length for digital power levels of 0.1 to 1 mW. For comparison, the total power per channel of the ATLAS SCT ABCD chip is about 3 mW and the digital power consumption is 1.1 mW per channel at a 40 MHz clock frequency and  $V_{DD} = 4$  V [32]. The assumed analog power in Fig. 24 is 0.2 mW for 10 cm strip length, as obtained by a recent SiGe design that meets sLHC requirements [33]. The total power rises rapidly for strip lengths below 3 cm, dominated by the digital power.

It is essential to streamline the digital circuitry to reduce its contribution by analyzing the contributions of individual circuit blocks. Unfortunately, a common assumption is that the digital power is negligible, often exacerbated by simulations that underestimate digital power. Complex circuits often require simplifying the simulation by subdividing the system into circuit blocks that don't include the actual circuitry, but provide the proper response to various actuations. This can be done, but the characteristics of the circuit blocks must be assessed correctly (and should be

<sup>1</sup> The above expression is often derived from the energy stored in a capacitor. However, this energy will be returned when the capacitor is discharged, so after the leading and trailing edges of a pulse the net energy is zero. In reality, the power is dissipated by the charge and discharge current flow in the circuit's series resistance (see Ref. [1]). This is one of many examples where the wrong physics yields a correct result – until one digs deeper.



**Fig. 24.** Total power in kW per  $\text{m}^2$  vs. strip length for various levels of digital power per channel  $P_{\text{dig}}$ .

experimentally verified) for all operating conditions. Simulations don't always give the right answers.

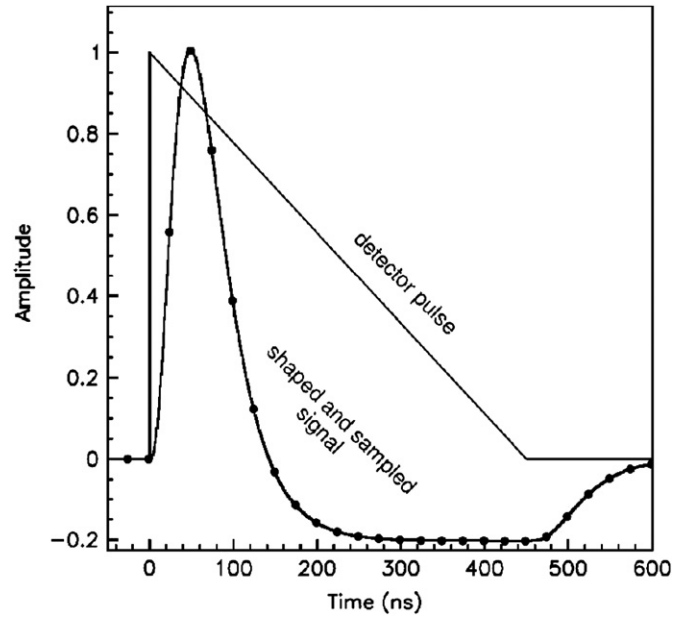
#### 6.4. Current sensitive front-ends (ATLAS liquid Ar calorimeter)

Most detectors designed for high energy resolution utilize charge-sensitive preamplifiers, but the ATLAS LAr calorimeter is subject to constraints that require a special front-end [34,35]. Carrier mobility in liquid Ar is rather small, so the individual cells in the ATLAS EM calorimeter have small electrode spacings to allow high event rates. Thus the rise time is a few ns and the pulse then decays over the maximum drift time in the LAr gaps of about 400 ns. This corresponds to a current of about 2–3  $\mu\text{A}/\text{GeV}$ , so the preamplifier must operate at input currents in the range of nA to mA. The calorimeter requires rather long cables connecting the electrodes to the cryostat feedthrough, so in the barrel the preamp is a current-sensitive transimpedance amplifier with an input impedance of 50  $\Omega$  to match the transmission line impedance. The amplifier gain is 3000 V/A. The forward systems have a higher detector capacitance, so the impedance is 25  $\Omega$  and the gain is 500–1000.

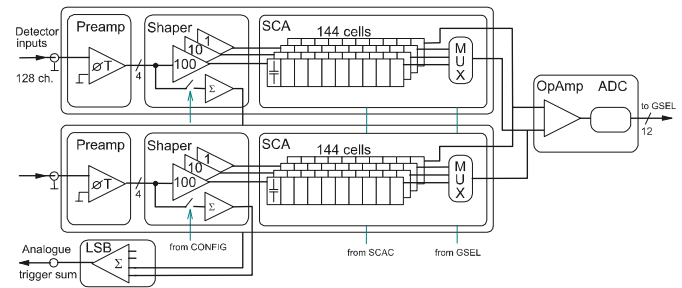
As in many cases, alternative solutions can also provide acceptable results. The hadronic endcap (HEC) utilizes cryogenic preamps mounted directly in LAr.

To avoid pileup a shorter pulse is required than the detector's 400 ns collection time, so a CR-(RC)<sup>2</sup> shaper with a 13 ns time constant is used. The CR “differentiator” reduces the pulse width and the two “integrators” reduce the bandwidth, which reduces the voltage noise. This degrades the electronic noise by only about 10% relative to an “ideal” filter. The convolution of the detector pulse and the shaping filter yields a peaking time of about 50 ns and a 20% undershoot that extends to about 600 ns. Fig. 25 shows the detector current pulse and the shaper output.

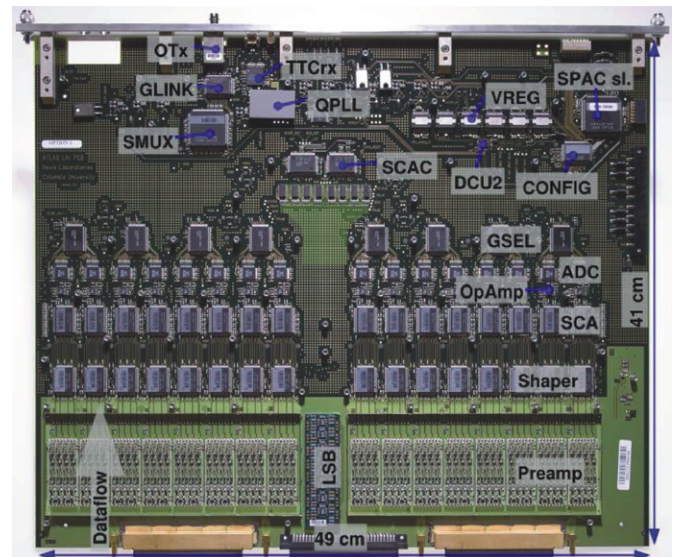
The preamp output is fed to three paralleled shapers with different gains to span the required dynamic range (Fig. 26). The gain ratio of a higher-to-lower gain is 10. The output signals are then sampled at 40 MHz by 144-cell switched capacitor arrays, which record the analog signals and store them during the L1 trigger latency. The preamplifiers use discrete components and the switched capacitor arrays combine 16 pipeline channels in custom ICs using the DMILL process. The remaining circuitry uses commercial ICs. Fig. 27 shows the combined analog-digital board.



**Fig. 25.** Detector current pulse and the shaper output of the ATLAS LAr front-end electronics [35].



**Fig. 26.** Block diagram of the ATLAS LAr calorimeter front-end electronics [35].



**Fig. 27.** Top layer of the ATLAS LAr front-end board combining the low-noise front-end electronics with digital readout circuitry [35].

## 7. Radiation resistance

### 7.1. Sensors

As noted in the discussion of electronic noise, displacement damage in the sensor leads to an increase in leakage current

$$I_R = I_0 + \alpha \Phi A d \quad (11)$$

which in the electronics increases the shot noise

$$Q_{ni}^2 = 2eI_d F_i T_s. \quad (12)$$

The leakage current drops exponentially with temperature

$$I_R(T) \propto T^2 e^{-E/2kT} \quad (13)$$

so even moderate reductions in temperature bring a significant improvement.

The electronic shot noise can be reduced by choosing short shaping times. Furthermore, reducing the area of a detector element reduces the leakage current per channel, so the shot noise is also smaller. Again, segmentation is advantageous.

The second major effect of displacement damage is an increase in the required operating voltage because of space charge buildup. The formation of defects that capture electrons leads to a relatively uniform distribution of negative space charge in the active region. At a certain fluence in an n-type substrate the negative space charge buildup compensates the positive space charge from the dopant core nuclei in the depleted region, so a reduced bias voltage will yield full charge collection. However, at high damage levels, typically at fluences beyond the  $10^{13} \text{ cm}^{-2}$  range, the damage space charge dominates. Without applying sufficient voltage to override the field due to the space charge, charge collection efficiency is reduced. Note that the bias polarity is no longer set by the original doping, since the force of the fixed charge can be overcome by an electric field in either direction, but in practice the pre-radiation polarity is maintained. The space charge buildup requires that detectors that initially may have had full charge collection with tens of volts need bias levels of hundreds of volts.

The electrical activity of the radiation defects is subject to both annealing and anti-annealing, where the latter tends to dominate and increases with temperature. As a result, silicon detectors are cooled to temperatures around  $-5^\circ\text{C}$ , which reduces the required bias voltage and also the leakage current. At high fluences as encountered in the pixel detectors, the bias voltages can be maintained at reliable levels by thinning the sensor. Furthermore, operation can continue below full charge collection by using an n-type signal electrode, which attracts electrons that will induce the most signal charge when moving near the signal electrode. This does reduce the signal and requires lower noise to maintain the required signal-to-noise ratio. Again, segmentation helps, as the decreased area of a detector element reduces the capacitance and the achievable noise level.

Cooling the semiconductor detectors is crucial for a further reason. The power dissipated in the sensor due to the leakage current together with the bias voltage leads to self-heating of the detector. Unless the sensor is cooled adequately, the increased power dissipation increases the temperature, which increases the bias current exponentially and leads to thermal runaway [25].

### 7.2. Electronics

In the electronic circuitry the key changes due to radiation damage are in the active devices. In MOSFETs charge buildup in the gate oxide changes the threshold voltage in digital systems and shifts the operating points in analog circuits. This is primarily due to ionization damage, i.e. charged particles or photons with

keV energies and greater. In digital circuitry this is especially critical, as it can change switching thresholds beyond a usable range. In analog circuitry operating points can be adjusted to compensate. For example, the noise level of the input stage is critically linked to the transconductance (output signal current vs. input voltage), which depends on the operating current. In many circuits the degradation in noise with radiation damage is due to the reduction in operating current. If the circuit is configured to shift the gate voltage to maintain the operating current, the radiation resistance is greatly improved (see Ref. [1]). In bipolar transistors the increase in base current with displacement damage also adds to the noise.

Digital circuitry is most susceptible to radiation damage. The major problem is the shift in switching thresholds, which initially leads to random switching and then to complete failure. Single event upset is another problem, where the triggering of a single transistor can lock up a digital circuit. The small feature sizes in the digital circuitry of modern ICs exacerbate this problem, as the charge required to change the gate voltage is very small, so a fraction of the energy deposition from a single particle may be enough.

The probability of single event upset can be reduced by special radiation hard IC fabrication processes, as promoted by the military. However, single event upset is not critical in all digital stages, so in conventional commercial processes additional radiation resistance can be achieved by carefully analyzing the digital circuitry and implementing redundancy at critical stages, i.e. requiring that two or more parallel circuits give the same result. The probability of both being subject to simultaneous single event upset is much smaller. An example of this scheme is described in Section 9.1.

Overall, the ability to achieve high radiation resistance with standard commercial processes is a major contribution to detector electronics in high-luminosity experiments. Unlike special radiation-resistant processes they are readily available and less expensive (although cost is still an issue). Furthermore, the processes used in large-scale commercial production are much better controlled and provide high yields.

## 8. Examples with extreme requirements I: silicon trackers

The LHC poses unprecedented challenges to detector designers. Work on suitable detector concepts began in the 1980s, culminating in final assembly in 2005–2007. A worldwide detector R&D program was necessary to develop the concepts and technologies, especially in the areas of sensors, microelectronics, and radiation effects. The results of this ongoing work also flowed into preceding experiments such as CDF, DØ, BaBar, and Belle. The next two sections are adapted from Ref. [1], which also discusses more underlying instrument physics.

Key LHC parameters include colliding proton beams of 7 TeV on 7 TeV to provide 14 TeV center-of-mass energy at a luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . For comparison, the SSC planned 10 TeV on 10 TeV; the higher energy allowed a ten-fold lower luminosity of  $10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  for essentially the same physics goals. The higher luminosity at the LHC increases the backgrounds and the radiation damage that the detectors must cope with.

The LHC bunch crossing frequency is 40 MHz with an average of about 20 interactions per bunch crossing and about 150 charged particles per unit of pseudo-rapidity  $\eta = -\log_{10}(\Theta/2)$ , where  $\Theta$  is the angle relative to the beam axis. Thus, the hit rate

$$n' = \frac{2 \times 10^9}{r_{\perp}^2} \quad (\text{cm}^{-2} \text{ s}^{-1}) \quad (14)$$



where  $r_{\perp}$  is the distance from beam axis. In a detector that subtends  $\pm 2.5$  units of rapidity, the total hit rate is  $3 \times 10^{10} \text{ s}^{-1}$ . At  $r_{\perp} = 14 \text{ cm}$  the rate is about  $10^7 \text{ s}^{-1} \text{ cm}^{-2}$ . This radial dependence is modified in the presence of a magnetic field.

### 8.1. Coping with high rates

As already noted above, the high hit rate can easily lead one to believe that very high speed electronics are needed. However, maintaining the required signal-to-noise ratio at fast shaping times drives up the required power. Segmentation, on the other hand, is much more efficient, as subdividing the detector into many small elements reduces the hit rate per element. For example, at  $r_{\perp} = 30 \text{ cm}$  the hit rate on a strip electrode of  $50 \mu\text{m}$  width and  $10 \text{ cm}$  length, i.e. an area of  $5 \times 10^{-2} \text{ cm}^2$ , is about  $10^5 \text{ s}^{-1}$ . This corresponds to an average time between hits of  $10 \mu\text{s}$ , so longer shaping times are allowable, which translates to lower power for given noise level. As explained in Ref. [1], with careful design the power requirements do not increase significantly in highly segmented detectors.

An additional problem is the large number of events per crossing. As the interaction region is spread out, vertex reconstruction with the appropriate  $z$ -resolution can resolve individual interactions. Again, segmentation helps. If a detector element is sufficiently small, the probability of two tracks striking it within one crossing is negligible. Tracks from different crossings can be separated if the electronics are capable of single-bunch time resolution, that is  $25 \text{ ns}$ , and time-stamped data are stored.

Semiconductor detectors are well-matched to these requirements and they are key components in all LHC experiments. Note that in the early days of the SSC, the widespread opinion was that semiconductor trackers could not be practical, even though sound arguments based on both physics and technology in support of these systems existed. Patterning the sensor at the “ $\mu\text{m}$ -scale” is straightforward and monolithically integrated electronics can be mounted locally with on-chip multiplexing or sparsification to reduce the cable plant. The fast collection times of semiconductor detectors are very advantageous; achieving collection times  $< 25 \text{ ns}$  with a  $300 \mu\text{m}$  thick sensor is quite practical.

### 8.2. Radiation damage

The high rates bring another problem with them; radiation damage, both in the sensors and the electronics. There are two sources of particles, beam collisions and neutron albedo from calorimeter, as was shown above in Fig. 7. Estimated fluences per year at design luminosity expressed in equivalent  $1 \text{ MeV}$  neutrons are  $5 \times 10^{13} \text{ cm}^{-2}$  at  $r = 10 \text{ cm}$  and  $2 \times 10^{13} \text{ cm}^{-2}$  at  $r = 30 \text{ cm}$ . The corresponding ionizing doses are  $30 \text{ kGy}$  ( $3 \text{ Mrad}$ ) at  $10 \text{ cm}$  and  $4 \text{ kGy}$  ( $400 \text{ krad}$ ) at  $30 \text{ cm}$ . In reality, complex maps are required of the radiation flux, which is dependent on local material distribution (see examples in the ATLAS TDR [12]).

As discussed in Section 7, displacement damage in the sensor leads to an increase in leakage current. The leakage current drops exponentially with temperature, so even moderate reductions in temperature bring a significant improvement. Cooling is also crucial to avoid thermal runaway due to self-heating of the detector [25]. The electronic shot noise can be reduced by choosing short shaping times. Furthermore, reducing the area of a detector element reduces the leakage current per channel, so the shot noise is also smaller.

Segmentation also reduces the input load capacitance to the front-end, so lower noise is possible, which allows a larger signal loss with radiation damage. As explained above, this allows pixel detectors to sustain high fluences close to the interaction region.

### 8.3. Layout

General purpose detectors provide full coverage by a combination of barrel and disk layers. As described above, both ATLAS and CMS use barrels in the central region and disks in the forward regions to provide the required coverage and tracking performance with minimum silicon area. The ATLAS SemiConductor Tracker (SCT) has about  $60 \text{ m}^2$  of silicon with  $6 \times 10^6$  strip detector channels, augmented by a gaseous outer tracking detector. After going rather far in the development of a mixed silicon-gaseous detector system, CMS decided to build an all-silicon tracker with about  $230 \text{ m}^2$  of silicon and  $10^7$  strip detector channels. Both ATLAS and CMS use pixel devices covering  $1\text{--}2 \text{ m}^2$  with  $50\text{--}100$  million channels at the inner radii ( $< 15 \text{ cm}$ ) because of their superior pattern recognition at high track densities and high radiation resistance. The small capacitance allows low noise  $Q_n \approx 200e$  at sufficiently fast shaping times, so the system starts out with a very high signal-to-noise ratio. These performance reserves allow greater degradation of signal and noise with radiation damage than in a silicon strip system, which extends the lifetime. Strips take over at larger radii to minimize material and cost.

In ATLAS four layers of silicon strips are used at radii of  $30, 37, 44$ , and  $51 \text{ cm}$  inside a superconducting solenoid with a  $2 \text{ T}$  magnetic field (Fig. 8). Beyond  $56 \text{ cm}$  radius a 70-layer straw-tube gaseous tracking and transition radiation detector (TRT) provides at least 40 hits per track. The TRT is operating close to its limits at  $10^{34}$  luminosity, but was retained for cost budget reasons. At both ends an array of 9 disks arranged at distances from  $|z| = 0.85$  to  $2.7 \text{ m}$  provide coverage at rapidity  $|\eta| > 1.2$ . Resolution in the barrel is determined by the strip pitch of  $80 \mu\text{m}$  in  $r\phi$  and  $40 \text{ mrad}$  small angle stereo in  $z$ . In the disks the strips go radially, with pitches ranging from  $55$  to  $90 \mu\text{m}$ , again using a  $40 \text{ mrad}$  stereo-angle for  $r$ -resolution. All modules are double-sided, formed by gluing two single-sided detectors back-to-back under an angle of  $40 \text{ mrad}$ .

Three layers of pixel detectors at  $5.1, 9.9$ , and  $12.3 \text{ cm}$  in the central region and three pixel disk layers at each end provide the two track resolution and radiation resistance required close to the interaction region. The pixels are  $50 \mu\text{m} \times 400 \mu\text{m}$ , with the long dimension along the beam axis to accommodate inclined tracks. The pixel subsystem is enclosed within the pixel support tube and can be inserted or removed as a unit to facilitate replacement.

Cooling both the sensors and the readout electronics is critical in these systems. Services providing both power and cooling are brought out to patch panels at the cryostat wall through the gaps between the TRT subunits, which leads to significant local increases in the material distribution.

Estimated fluences after 10 years of operation are estimated to be  $10^{15} \text{ cm}^{-2}$  ( $1 \text{ MeV}$  neutron equivalent) with a total dose of  $50 \text{ Mrad}$  at the innermost pixel layer and a fluence  $2 \times 10^{14} \text{ cm}^{-2}$  at the inner strip layer. The  $p^+$ -on- $n$  sensors used in the SCT must be biased for full charge collection. Space charge buildup and anti-annealing increase the required detector bias voltage to  $> 350 \text{ V}$ , so the sensors have been designed to sustain  $500 \text{ V}$ . In production about 10% of the modules exhibited the onset of high bias current well below the  $500 \text{ V}$  required in the acceptance tests [17]. Most of these modules met specifications after operating them with gradually increasing bias voltage over several hours. Both the traditional  $\langle 111 \rangle$  and  $\langle 100 \rangle$  orientations were studied and  $\langle 111 \rangle$  chosen because of easier availability. The pixel modules use oxygenated sensors, which allow full voltage operation over  $> 5$  years [36]. Unlike the strip sensors, which are  $p^+$ -on- $n$ , the pixel sensors are  $n^+$ -on- $n$ , which are also usable at voltages below full collection. As a consequence, detector performance deteriorates gradually with radiation damage as the signal-to-noise ratio falls off.

In both the disks and the barrels, the modules are “shingled” so that adjacent modules overlap circumferentially to provide full coverage and facilitate relative position calibration (Fig. 28). The cant angle of the detectors is chosen to minimize the resolution spread due to Lorentz deflection of the carriers in the 2 T magnetic field [37,38]. The resolution in  $r\phi$  is 12  $\mu\text{m}$  in the pixel and 16  $\mu\text{m}$  in the strip system. The respective resolutions in  $z$  are 66 and 580  $\mu\text{m}$ . The ATLAS SCT uses the same strip detector design throughout the detector for budgetary reasons and to minimize the number of module designs, which allows more efficient fabrication and testing.

As noted above, CMS uses an all-silicon tracker with 2.4 m diameter and 5.4 m length in a 4 T solenoidal magnetic field. Fig. 13 shows the layout. Strip detectors are used in all layers except at the smallest radii, where the interaction region is surrounded by two barrel layers of pixel detectors at 4 and 7 cm for low luminosity running and at 7 and 11 cm at high luminosity. Two endcap pixel disks cover radii from 6 to 15 cm. Strip pitches range from 80 to 205  $\mu\text{m}$  and the pixel size is 100  $\mu\text{m} \times 150 \mu\text{m}$  [43].

CMS uses both single- and double-sided strip detectors. Double-sided detectors are used in layers 1, 2, 5, and 6 of the barrel and in rings 1, 2, and 5 of the disks. As in ATLAS the double-sided modules use two single-sided sensors, glued back-to-back to form a small stereo-angle. CMS uses a somewhat larger stereo-angle of 100 mrad. The endcap disks consist of wedge shaped segments, each covering 1/16 of  $2\pi$ .

In total the CMS tracker implements 25 000 silicon strip sensors covering an area of 210  $\text{m}^2$ . The  $9.6 \times 10^6$  strips are read out by 75 000 readout ICs. About 25 million wire bonds interconnect strip sensors and readout ICs. Detector segmentation is

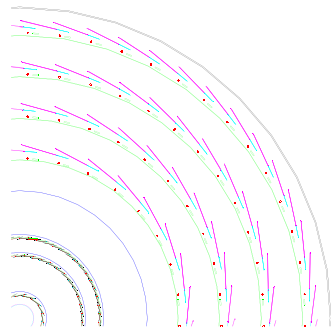


Fig. 28. Axial quarter view of the ATLAS semiconductor tracker illustrating shingling of detector modules to provide overlap and compensate for Lorentz deflection of the collected charge (ATLAS TDR [12]).

chosen so that typical channel occupancies are about 1% throughout the detector. Position resolution in  $r\phi$  is strongly affected by the approximately  $30^\circ$  Lorentz angle of the electron drift in the 4 T magnetic field. The barrel strip detectors are tilted by  $9^\circ$  to compensate. The barrel pixel geometry is deliberately chosen so that this large Lorentz angle induces significant sharing of charge across neighboring cells. This yields spatial resolutions  $\sigma_{r\phi} \approx 10 \mu\text{m}$  and  $\sigma_z \approx 15 \mu\text{m}$ .

In the inner layers the strip length is about 10 cm, whereas in the outer region the strip length is doubled, which increases the electronic noise. To compensate, the signal is increased by using 500  $\mu\text{m}$  thick sensors instead of the normal 320  $\mu\text{m}$  devices used elsewhere. Utilization of 150 mm diameter wafers that became available at the appropriate time provided the cost savings needed for this huge silicon system.

Sensors use the  $\langle 100 \rangle$  orientation to minimize surface damage. This yields a somewhat lower interstrip capacitance after irradiation [54]. The strips are AC-coupled with integrated polysilicon resistors. The inner region utilizes lower resistivity material (1.25–3.25  $\text{k}\Omega \text{cm}$ ) to delay space-charge inversion, whereas the 500  $\mu\text{m}$  thick sensors in the outer layers use 3.5–7.5  $\text{k}\Omega \text{cm}$  material [55,56]. The lower resistivity sensors start with a higher depletion voltage but end with a lower operating voltage after space charge inversion and 10 years of LHC operation. The strip pitch is 80–183  $\mu\text{m}$  in the barrel and up to 205  $\mu\text{m}$  in the disks with no intermediate strips.

#### 8.4. Readout electronics

Both ATLAS and CMS distribute the total number of tracks over many detector segments to reduce the rate per channel and reduce the double-hit probability. For example, in ATLAS the occupancy in the pixel system is  $4.4 \times 10^{-4}$  at 4 cm radius and  $6 \times 10^{-5}$  at 11 cm radius. In the strip system the occupancies are  $6 \times 10^{-3}$  at 30 cm and  $3.4 \times 10^{-3}$  at 52 cm radius. Thus, the choice of shaping time is not driven by rate considerations, but by the requirement for 25 ns single-bunch time resolution.

##### 8.4.1. CMS readout electronics

The CMS readout is a direct descendant of the systems used at LEP and utilizes full CMOS circuitry that exploits switched capacitor techniques. Fig. 29 shows the block diagram of the readout IC, the APV25 [42]. Each strip is read out by a charge-sensitive amplifier followed by a switchable unity gain inverter to allow both p- and n-strip readouts. Subsequently, a 50 ns CR-RC shaper drives a 192-stage analog pipeline to accommodate up to 4  $\mu\text{s}$  trigger latency. On receipt of a trigger a switched capacitor analog

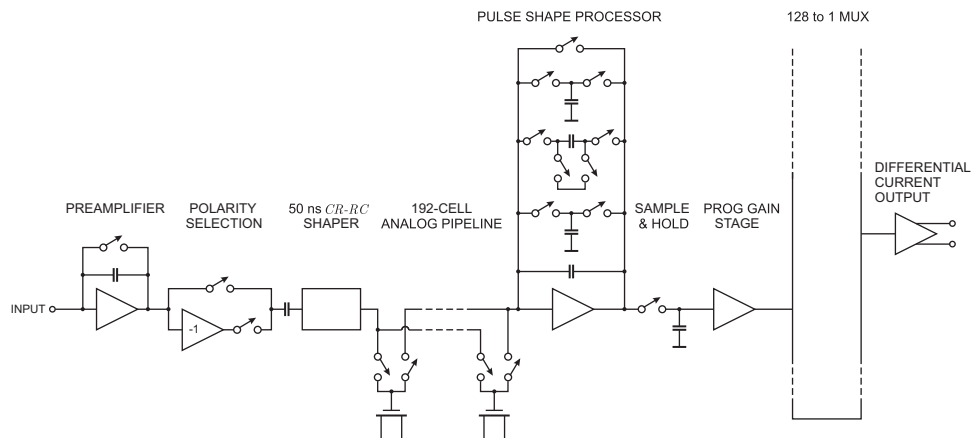


Fig. 29. Block diagram of the APV25 readout IC used by CMS.



pulse processor applies a weighted sum algorithm to provide the desired single-bunch time resolution.

Development of the pulse processor was originally motivated by the lack of sufficiently fast CMOS processes to allow efficient operation at the shaping times required for single-bunch time resolution [40]. The underlying notion is to use rather slow pulse shaping and then to apply a deconvolution algorithm to reconstruct the fast components of the input signal. For the sampled output of a CR-RC shaper with the step response

$$v(t) = \frac{t}{\tau} e^{-t/\tau} \quad (15)$$

this can be implemented by forming the weighted sum of three successive samples

$$V_k = w_1 V_k + w_2 V_{k-1} + w_3 V_{k-2} \quad (16)$$

with the weights [39,40]

$$w_1 = \frac{1}{x} e^{x-1}, \quad w_2 = -\frac{2}{x} e^{-1}, \quad w_3 = \frac{1}{x} e^{-(x+1)}. \quad (17)$$

The weights depend on the sampling interval normalized to the shaping time constant  $x = \Delta t / \tau$ . For a step input the result of the deconvolution is zero. However, for a finite rise time the result is a short pulse with the duration of the rise time. The APV25 uses a time constant  $\tau = 50$  ns in the CR-RC filter and samples the output at 40 MHz, so  $x = 0.5$  and the weighting factors:

$$w_1 = 1.2, \quad w_2 = -1.5, \quad w_3 = 0.45.$$

Fig. 30 shows the APV25 output in peak and deconvolution mode. As this is a crude form of differentiation, the signal is reduced and the noise bandwidth increased, so the noise in deconvolution mode is higher than in peak mode. For example, the noise in peak mode of  $246e + 36e/\text{pF}$  increases to  $396e + 59.4e/\text{pF}$  in deconvolution mode. The latter is required for single-bunch timing resolution at the LHC, so that is the relevant noise in normal operation. Note that often the lower noise is quoted, as it is relevant for other applications.

Viewed somewhat differently, the CR-RC shaper is the anti-aliasing filter required before sampling the signal into the analog pipeline. The algorithm then produces a near triangular weighting function [40] with a peaking time of  $\Delta t$  and the noise indices  $F_i = 0.35$  and  $F_v = 1.84$ , similar to a conventional CR-RC<sup>3</sup> or CR-RC<sup>4</sup> filter.

The decision to utilize an all CMOS readout IC with the deconvolution processor was made early on and applied in a series of designs using different fabrication processes. Fortunately, the demonstration of excellent radiation resistance in standard commercial “deep sub-micron” CMOS opened the path to an efficient implementation in  $0.25 \mu\text{m}$  CMOS [42]. The APV25 chip combines 128 readout channels with an output multiplexer and is 7.1 mm wide and 8.1 mm long [41].

The analog output signals are transmitted to the off-detector electronics through optical links using edge-emitting semiconductor lasers operating at a standard telecommunications wavelength of 1310 nm. Off-detector, the optical signals are received by a photodiode-amplifier on the “front end driver”, which digitizes and processes the signals, subtracts pedestals and stores the results in a local memory. When operating at the maximum trigger rate, cluster finding is applied to reduce the data volume.

#### 8.4.2. ATLAS readout electronics

ATLAS chose a readout system that sought to efficiently balance the technology against cost, while meeting the physics requirements. The goal in these large systems is not to provide the best possible performance, but to maintain adequate performance over the lifetime of the detector. After a lengthy process of

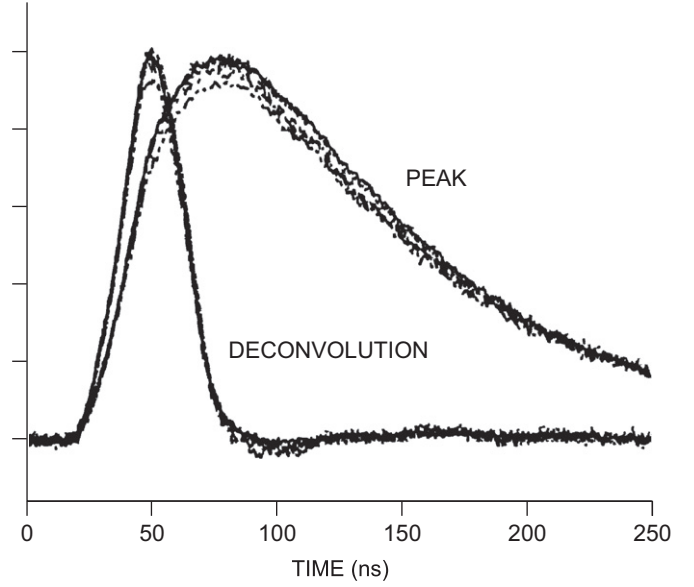


Fig. 30. APV25 pulse shapes in peak and deconvolution mode, shown for input capacitive loads ranging from 2 to 20.5 pF. The reduced pulse height at 17.5 and 20.5 pF indicates that the preamplifier bandwidth is becoming marginal. (Adapted from Ref. [41].)

testing and comparing options, the collaboration concluded that these goals can be achieved by a binary readout, i.e. a system that just records the presence of a hit, so the output only provides a time stamp with a series of hit addresses. This technique also lends itself readily to on-chip zero-suppression, which reduces the cost and space requirements of the readout lines. Threshold scans allow pulse height and noise measurements for diagnostics, but this is only necessary infrequently to monitor changes in response, e.g. due to radiation damage.

Fig. 31 shows the block diagram of the readout IC. The front-end utilizes time-invariant filtering. The bandwidths of the cascaded amplifiers needed to provide the necessary gain are tailored to provide approximately a CR-RC<sup>3</sup> response. A comparator fires when a signal exceeds threshold and the time is stored in a digital pipeline, whose length accommodates the ATLAS Level 1 trigger latency. Data sparsification and compression circuitry reduce the data volume that must be read out. More details follow below.

#### 8.4.3. Required signal-to-noise ratio in a binary readout system

In binary readout systems the threshold must be set low enough to capture the desired portion of the amplitude spectrum, but not so low that the rate of noise pulses is too high.

The signal for minimum ionizing particles is a Landau distribution, where for 99% efficiency in a  $300 \mu\text{m}$  thick detector the threshold must be set to about one-half the of the most probable charge  $Q_0$ . Assume that the minimum signal to be measured is  $f_L Q_0$ . Tracks passing between two strips will deposit charge on both strips. The ability to distinguish one-hit from two-hit clusters improves the obtainable position resolution, as two-hit clusters are assigned mid-way between two strips. If the fraction of the signal to be detected is  $f_{sh}$ , the circuit must be sensitive to signals as low as

$$Q_{\min} = f_{sh} f_L Q_0. \quad (18)$$

The required threshold-to-noise ratio for a noise occupancy  $P_n$  in a time interval  $\Delta t$  is

$$\frac{Q_T}{Q_n} = \sqrt{-2 \log \left( 4 \sqrt{3} T_s \frac{P_n}{\Delta t} \right)}. \quad (19)$$

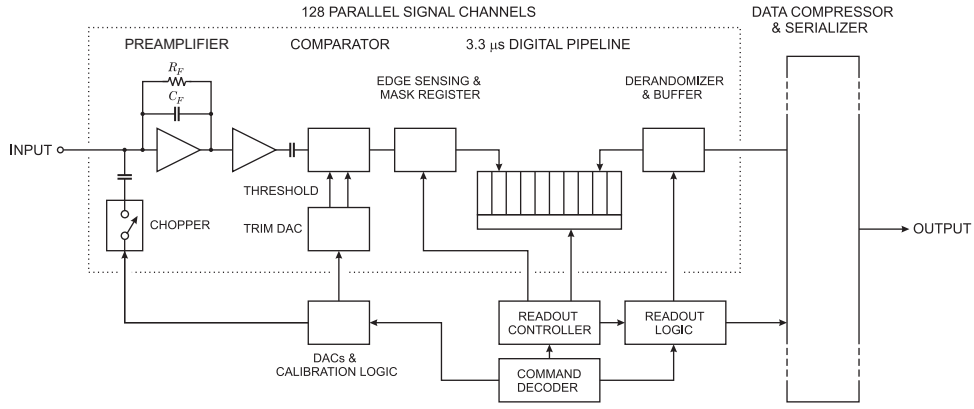


Fig. 31. Block diagram of the ABCD readout IC for the ATLAS SCT.

In the strip system the average hit occupancy is about  $5 \times 10^{-3}$  in a time interval of 25 ns. If we allow a noise occupancy of  $10^{-3}$  at a shaping time of 20 ns, this corresponds to  $Q_T/Q_n = 3.2$ .

The threshold uniformity is not perfect. The relevant measure is the threshold uniformity referred to the noise level. For a threshold variation  $\Delta Q_T$ , the required threshold-to-noise ratio becomes

$$\frac{Q_T}{Q_n} = \sqrt{-2 \log \left( 4 \sqrt{3} T_s \frac{P_n}{\Delta t} \right)} + \frac{\Delta Q_T}{Q_n}. \quad (20)$$

If  $\Delta Q_T/Q_n = 0.5$ , the required threshold-to-noise ratio becomes  $Q_T/Q_n = 3.7$ . To maintain good timing, the signal must be above threshold by at least  $Q_n$ , so  $Q_T/Q_n > 4.7$ .

Combining the conditions for the threshold

$$\left( \frac{Q_T}{Q_n} \right)_{\min} Q_n \leq Q_{\min} \quad (21)$$

and signal (Eq. (18))

$$Q_{\min} = f_{sh} f_L Q_0 \quad (22)$$

yields the required noise level

$$Q_n \leq \frac{f_{sh} f_L Q_0}{(Q_T/Q_n)_{\min}}. \quad (23)$$

If charge sharing is negligible  $f_{sh} = 1$ , so with  $f_L = 0.5$ ,  $Q_0 = 3.5$  fC, and  $(Q_T/Q_n)_{\min} = 4.7$ , the required noise level  $Q_n \leq 0.37$  fC or  $Q_n \leq 2300e$ . If the system is to operate with optimum position resolution, i.e. equal probability of one- and two-hit clusters, then  $f_{sh} = 0.5$  and  $Q_n \leq 0.19$  fC or  $Q_n \leq 1150e$ . ATLAS requires  $Q_n \leq 1500e$ .

#### 8.4.4. ATLAS SCT readout implementation

ATLAS adopted a bipolar transistor front-end with CMOS digital circuitry. Initial prototypes used separate bipolar and CMOS ICs. The production device utilizes the DMILL BiCMOS process that combines all of the circuitry in a single chip, the ABCD chip [52,32]. Each chip includes 128 channels, on a  $6.4 \times 4.5$  mm<sup>2</sup> die, bondable to a 50 μm pitch. Pitch adapters make the transition to the detector strip pitch of 80 μm. Designing the ICs for the smaller pitch provides space between adjacent ICs for bypass capacitors and wire bonds as shown below. The analog portion uses continuous shaping, approximating a CR-RC<sup>3</sup> response with a peaking time of 20 ns. At the nominal operating threshold of 1 fC this yields a time walk of 12 ns for signals of 1.2–10 fC. The double-pulse resolution for two successive 4 fC pulses is 50 ns. The operating current of the input transistor

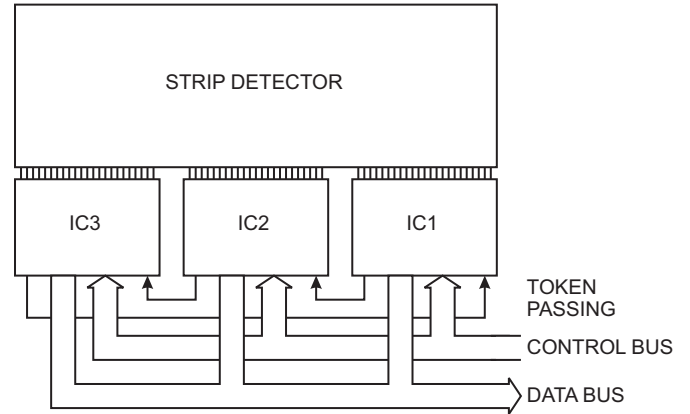


Fig. 32. Multiple ICs are ganged to read out a strip detector. The right-most chip IC1 is the master. A command on the control bus initiates the readout. When IC1 has written all of its data it passes the token to IC2. When IC2 has finished it passes the token to IC3, which in turn returns the token to the master IC1.

is adjustable to optimize noise with radiation damage and the total power is 1.3–1.8 mW/ch.

On-chip DACs control the threshold and operating point. Trim DACs on each channel fine tune the thresholds to compensate for threshold nonuniformity from channel to channel, bringing the threshold dispersion well below the noise level. This technique is even more important when power and readout area are minimized, as in pixel devices, so the efficacy of trimming will be illustrated in Section 9.

Each chip also includes digitally controlled calibration circuitry with a DAC-controlled injection level, shown in Fig. 31. All control and output signals are digital and each module communicates with the off-detector electronics through optical fibers. The serial link and token passing (see Fig. 32) present single-point failure modes, so redundant readout modes are incorporated. In token passing defective chips within a module can be bypassed. Normally, each module has two readout lines, one for each side. Should the master of one side fail, the other side's master takes over and both sides are read out through one line. For a more detailed description of the ABCD readout IC see Ref. [32].

The adopted IC fabrication process was specially designed for LHC applications and lacked a strong commercial base. Consequently, process control was not fully developed and the overall yield was about 20%. The project required about  $5 \times 10^4$  ICs, so this required a fast testing system, which was custom designed to provide the necessary throughput to match the production schedule [44].

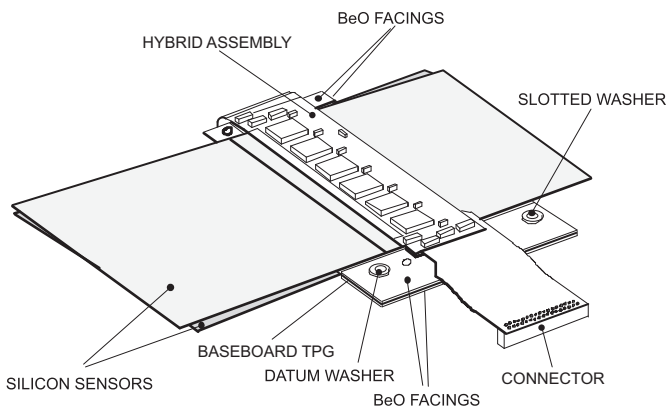
### 8.5. Detector modules

CMS uses a conventional module configuration with ceramic hybrids connected at the ends of the detectors. ATLAS adopted a novel module design, so it will be discussed in more detail. Fig. 33 shows the module layout. Connecting the electronics at the mid-point of the barrel modules reduces the noise contribution of the strip resistance. The electronics hybrid uses a four-layer polyimide substrate shown in Fig. 34 [45,46]. This reduces material and also allows the electronics for both sides to be placed on the same layer, as the hybrid can be wrapped around the detector (Fig. 35). A close up of the ICs mounted on the hybrid is shown in Fig. 36, which also shows the space between adjacent ICs for bypass capacitors and wire bonds.

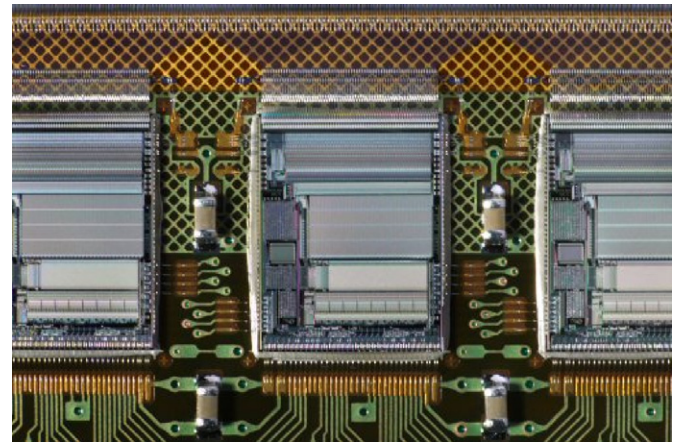
Two single-sided p-on-n sensors are glued back to back to form a 40 mrad stereo-angle. Each sensor has 784 strips on an 80  $\mu\text{m}$  pitch. Two sensors are butted to provide an overall strip length of 126 mm. An intermediate baseboard of thermalized pyrolytic graphite (TPG) provides support and a high-conductivity cooling path to the mounting stave of the barrel support

structure. The hybrid is mounted on spacers to prevent direct heat transfer from the readout ICs to the sensors. This bridge configuration and the high thermal conductivity of the TPG heat spreader between the sensors ensures that the sensors are cooled sufficiently to limit anti-annealing and thermal runaway after radiation damage. Results of a finite element thermal simulation are shown in Fig. 37 [45]. The disk modules use a similar structure, but the electronics are end mounted [47–49]. Expressed in percent of a radiation length  $X_0$ , the material in an SCT module consists of silicon sensors and adhesive: 0.612, baseboard and BeO facings: 0.194, ICs and adhesive: 0.063, Cu/polyimide hybrid: 0.221, passive components: 0.076, adding to a total of 1.17%  $X_0$ . At normal incidence the detector modules, the cooling and support structure, and the cabling in the four-layer tracker add up to 0.1 $X_0$ .

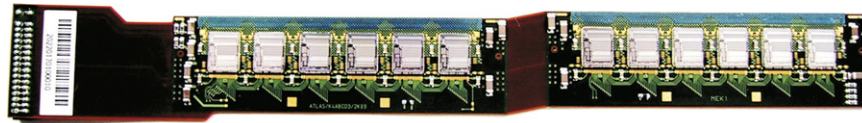
Electrical signals are transmitted by fully balanced LVDS links. Noise pickup from the digital electronics to the sensors is negligible and modules operated in systems mock-ups show



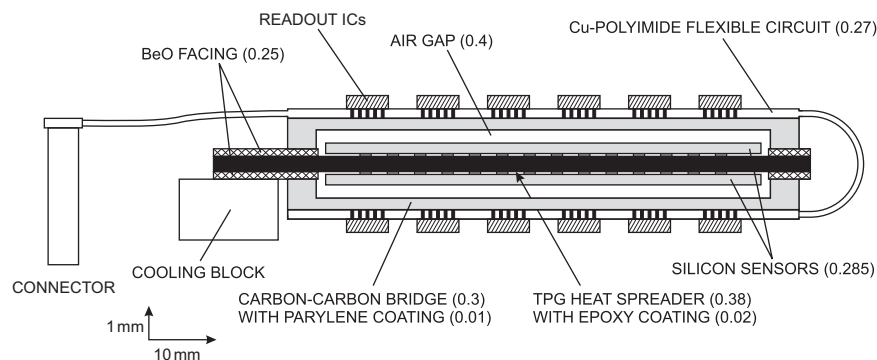
**Fig. 33.** ATLAS SCT barrel detector module. Two single-sided sensors are glued back-to-back with an intermediate TPG heat spreader. The “ear” extending from the module attaches to the support/cooling stave of the SCT barrel structure. (Figure courtesy of T. Kondo.)



**Fig. 36.** Closeup of the SCT readout ICs mounted on the hybrid. The high-density wire bonds at the upper edges of the chips connect via pitch adapters to the 80  $\mu\text{m}$  pitch of the silicon strip detector at the upper edge of the photograph. The ground plane is patterned as a diamond grid to reduce material. (Photograph courtesy of A. Cioico.)

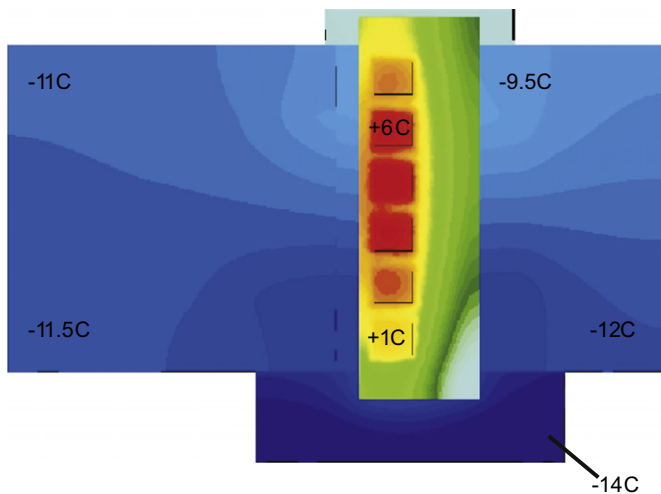


**Fig. 34.** SCT front-end ICs and associated components are mounted on a flex-hybrid that wraps around the module. The hybrid integrates electronics, interconnections and the connector for both sides. Bypass capacitors are visible adjacent to each IC and off the ends of the two arrays of readout ICs. (Photograph courtesy of T. Kondo.)

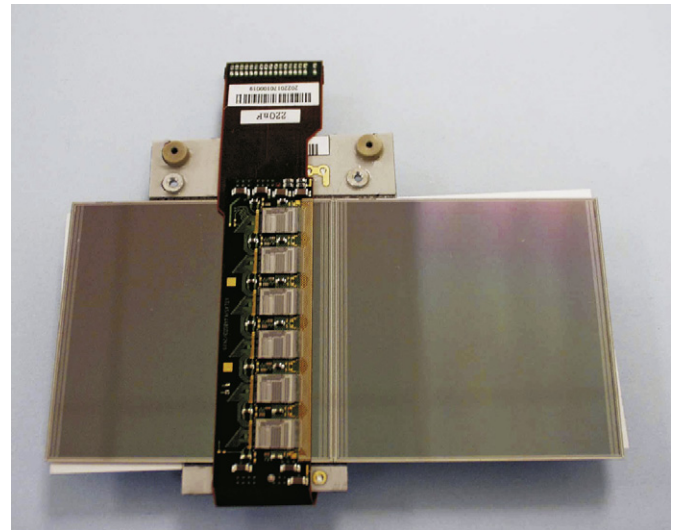


**Fig. 35.** Schematic cross-section of an SCT barrel module (vertical scale exaggerated). The hybrid is glued to a bridge to reduce heat transfer to the sensors. The height of the bridge still allows reliable wire bonding to the sensors. Thicknesses (in parenthesis) are in mm. (Figure courtesy of T. Kondo.)





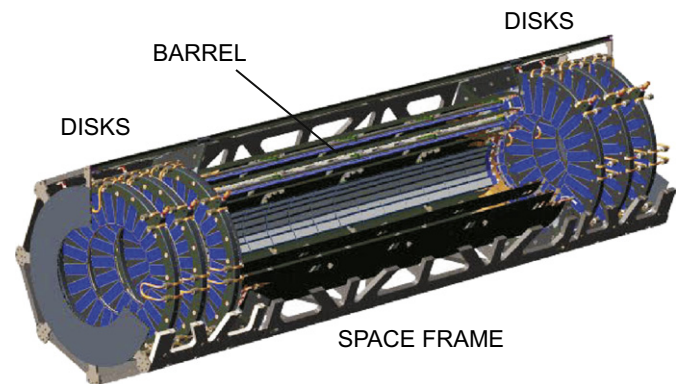
**Fig. 37.** Temperature distribution of an SCT detector module. (Figure courtesy of T. Kondo.)



**Fig. 38.** Photograph of an assembled SCT barrel module. The attached bar code allows component tracking during tracker assembly. (Figure courtesy of T. Kondo.)

negligible common mode noise and no cross-talk [50]. Fig. 38 shows a photograph of an assembled module. For a summary of test beam results see Ref. [51].

The scale of these projects does not allow the improvisation and last minute crash programs that characterize smaller projects. Small projects can make last minute changes and implement them rather quickly. Small systems also tend to be more accessible, so after some initial running it is common to take them out for rework. In huge detectors like ATLAS and CMS, removing the silicon systems is a major effort that necessitates significant downtime. Thus the reliability requirements are similar to systems in space. Both ATLAS and CMS have adopted extensive test and quality control procedures that track components through the production process and record test data at every step in a database [44,53,57]. Assembly and testing are distributed over multiple institutions [58,59], so uniform acceptance criteria must be established and enforced.



**Fig. 39.** The ATLAS pixel detector. The length of the detector is 1.4 m and the radius of the outermost pixel layer is 12 cm. (Figure courtesy of M.G.D. Gilchriese.)

## 9. Examples with extreme requirements II: Silicon pixel detectors

The high occupancy at the inner layers of semiconductor tracking detectors at high luminosity colliders precludes the use of strip detectors. Pixel detectors are key components of ATLAS, CMS [60,43], ALICE [61,62], and other systems. Unlike CCDs, these devices allow the selective readout of individual pixels, so they are often called random access pixel devices.

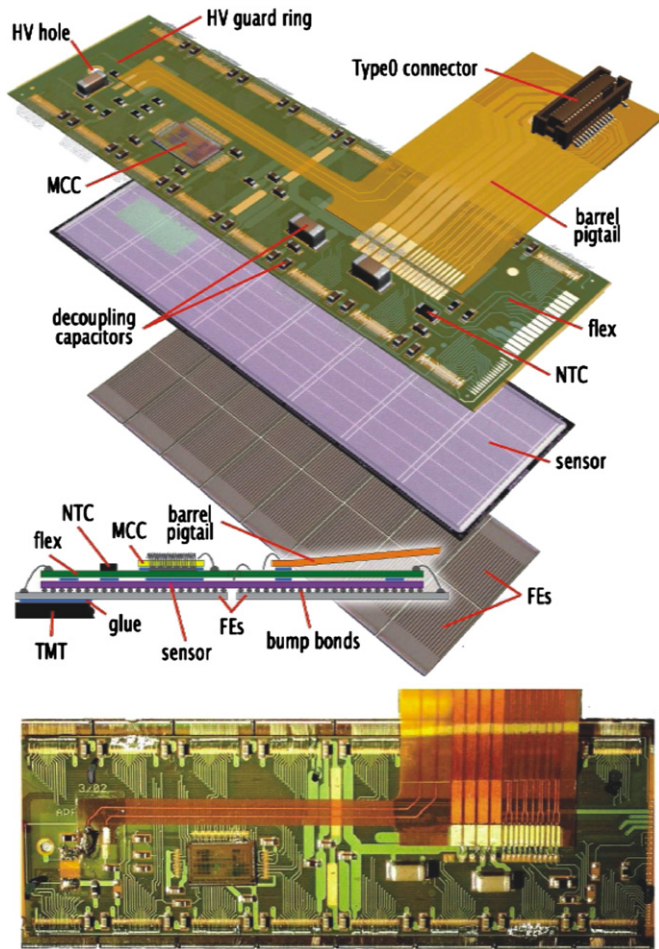
Small-scale two-dimensional segmentation allows pattern recognition at high track densities. The low capacitance provides a high signal-to-noise ratio, which allows degradation of both the sensor signal and electronic noise due to radiation damage. With the small detector elements the detector bias current per element is still small after radiation damage. The drawback is that the engineering complexity is at least an order of magnitude greater than in strip systems, so the path toward devices suitable for the LHC has been arduous. Along the way devices have been used successfully in DELPHI [63] and WA97 [64], but the LHC detectors are a significant step up in complexity and scale. See Ref. [65] for an overview of these early systems. Designs suitable for high-luminosity colliders began in the late 1980s [66–68] and have come to fruition for the LHC. The ATLAS pixel device will be described to illustrate the design techniques.

### 9.1. ATLAS pixel detector

Fig. 39 shows the layout of the ATLAS pixel detector [69]. The overall pixel detector has about  $2 \text{ m}^2$  of sensor area and  $10^8$  channels. The pixels are  $50 \times 400 \mu\text{m}^2$ , oriented along the beam axis to reduce distribution of the signal from inclined tracks over multiple pixels. Fig. 40 shows a barrel module, which consists of a  $6 \times 1.6 \text{ cm}^2$  silicon sensor wafer, onto which two rows of eight readout ICs with a total of 46 080 pixels are bonded by a two-dimensional array of solder bumps. A cross-section and more details of the bump bonding are shown in Fig. 41. For more details see Ref. [70] and for an overview of high-density interconnect technology see Ref. [71].

On a readout IC the pixels are arranged in 18 columns of 160 pixels. Each pixel cell contains a full electronic channel with control circuitry, described below. Tiling is accomplished without dead area by making the sensor pixels at the readout chip boundaries longer to bridge the gap. The electronic noise averages  $170e$ , obtained during simultaneous readout at a 40 MHz rate.

As shown in Fig. 40 the output pads of the readout ICs extend beyond the edge of the sensor to allow wire bonding to a flex-hybrid, which accommodates bypass capacitors, a readout controller IC, and power, control and readout bussing. Communication is through three links, input, output, and clock. The module

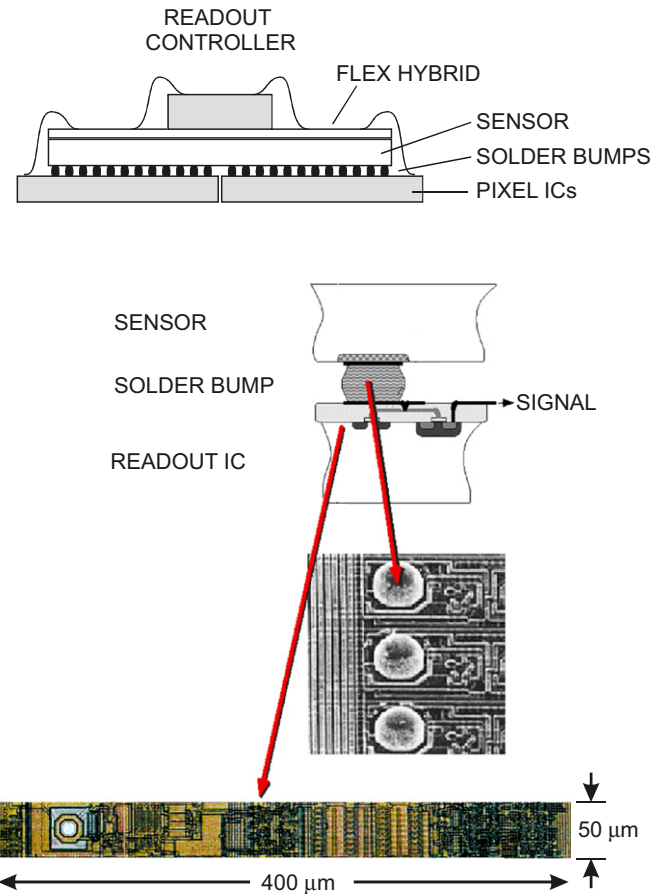


**Fig. 40.** ATLAS barrel pixel module. The silicon sensor provides the basis onto which the 16 readout chips are bump-bonded. The readout chips are wire bonded to a polyimide flex-hybrid, which provides the interconnections and the readout controller IC [16].

communicates through an optical package, up to 1 m distant, that connects to the off-detector electronics.

The sensors utilize oxygenated n-type silicon bulk with  $n^+$  electrodes [72,73]. Interelectrode isolation is provided by a contiguous “p-spray” [74]. The  $n^+$ -on- $n$  structure still provides good efficiency when operated below the voltage required for full charge collection, so it extends the overall detector lifetime after radiation damage. Sensors are 250  $\mu\text{m}$  thick. The pixels are direct coupled to the amplifiers, as space constraints do not allow the bias structures required per pixel with AC coupling. However, some form of common biasing is required for sensor testing and also to maintain a uniform potential distribution around a faulty bond. The ATLAS sensors implement a bias grid that provides punch-through biasing, which is inactive during operation. Tests indicate that the voltage required for full charge collection should remain below 600 V throughout ten years of LHC running [72]. Beyond this the signal degrades gradually, both due to incomplete charge collection and trapping, so the system still remains functional, albeit with reduced signal-to-noise ratio. The readout IC is fabricated in a standard “deep submicron” 0.25  $\mu\text{m}$  process and has shown only minor degradation after irradiation to 100 Mrad [76].

A block diagram of the circuitry in each cell is shown in Fig. 42. Each pixel cell contains a preamplifier, pulse shaping with a 30 ns rise time, a threshold comparator, a trim-DAC for pixel-by-pixel fine adjustment of the threshold, time stamp logic, and event buffering [77]. Current feedback in the charge-sensitive amplifier



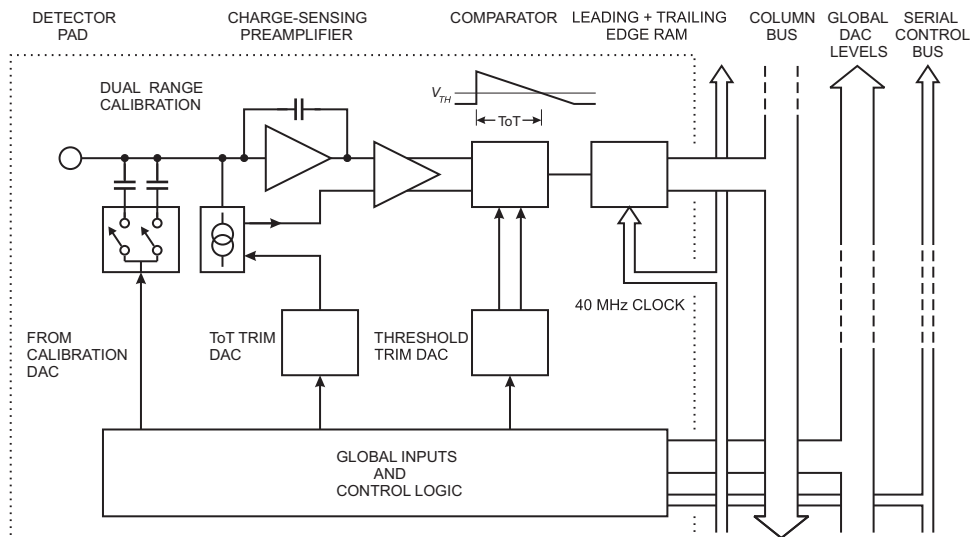
**Fig. 41.** Cross-section and bump bonding of an ATLAS pixel module. In the cross-section view the pixel ICs are at the bottom, bonded to the sensor above through a two-dimensional array of solder bumps. Connections from the pixel ICs to the flex-hybrid are by wire bonds. A pixel cell and bump bond are shown at the bottom. The bump bond pad is a 20  $\mu\text{m}$  diameter octagon with a 12  $\mu\text{m}$  opening in the passivation for the solder bump. (Figure courtesy of M. Garcia-Sciveres.)

provides a linear 500 ns–1  $\mu\text{s}$  discharge and also compensates for the sensor bias current. Pulse heights are digitized by measuring the time over threshold (ToT). Unlike the BaBar ATOM IC [75], this is designed to provide a linear response. The preamplifier is direct coupled to a two-stage differential amplifier. The reference level is generated in the preamplifier discharge block to track the baseline at the preamplifier output. The amplifier drives a differential comparator whose threshold is set by adjusting the baseline of the second gain stage. The PMOS input transistor with  $W = 25.2 \mu\text{m}$  and  $L = 0.6 \mu\text{m}$  operates at 8  $\mu\text{A}$  drain current. The feedback capacitor is 6 fF, yielding a 15 ns risetime with the sensor connected. Charge interpolation using the ToT yields a spatial resolution of 7–10  $\mu\text{m}$  in  $r\phi$ .

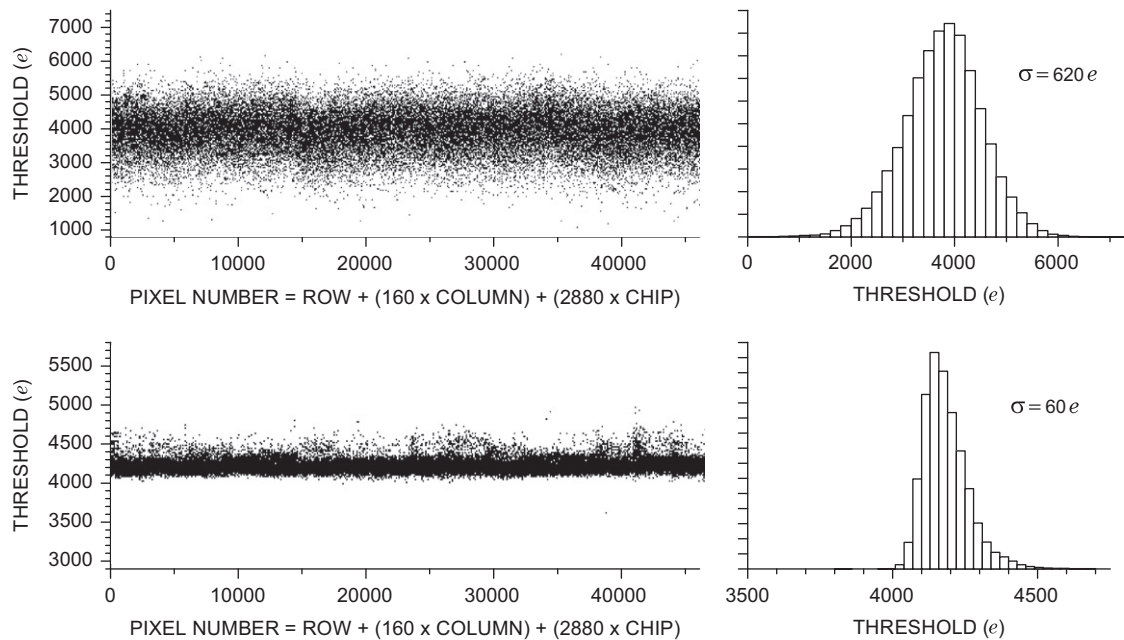
The small transistors required for this design do not provide adequate matching to keep threshold variations well below the noise level, so a fine adjustment is incorporated. Each pixel cell includes a seven-bit threshold trim DAC and a three-bit DAC to trim the ToT. Fig. 43 shows the threshold dispersion before and after trimming. Individual pixel cells can be selected for charge injection, masking the output of noisy pixels, or shutdown in case of cell failure.

Globally all critical bias currents and voltages on the chip are controlled by DACs (11 DACs with eight-bit resolution). A ten-bit DAC controls charge injection and another ten-bit DAC modifies the input current discharge to provide a measurement of the leakage current of each individual pixel. Two charge injection capacitors are incorporated in each pixel to provide a low range





**Fig. 42.** Block diagram of an ATLAS pixel cell. Each cell contains a full analog chain: a preamplifier, shaper, and threshold comparator. A buffer records time stamps of the leading and trailing edge of each pulse and immediately sends it to a buffer at the foot of each column.



**Fig. 43.** Threshold distribution of an ATLAS pixel module before and after trimming. A signal charge of 4000e is injected and the trim DACs in each pixel cell adjusted to minimize the threshold dispersion. (Data courtesy of M. Garcia-Sciveres.)

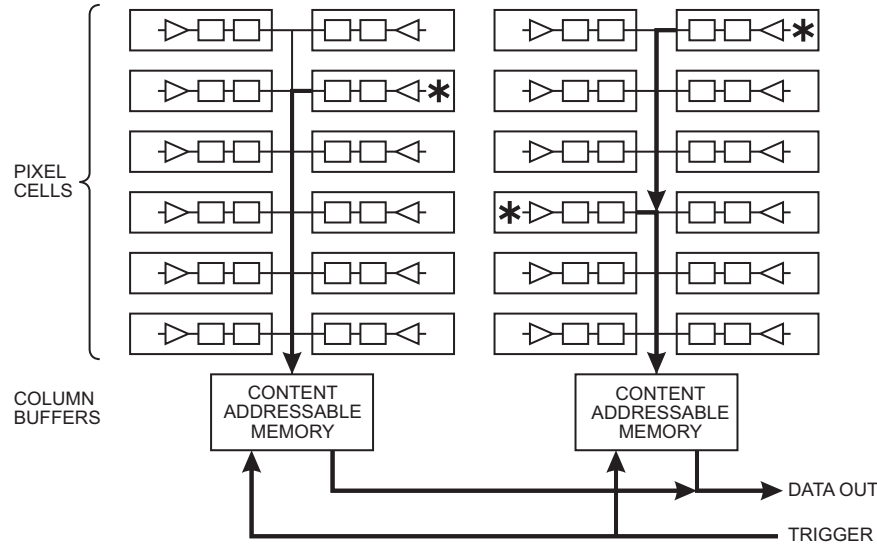
for noise and threshold measurements and a high range for calibration, time walk and cross-talk measurements.

The analog supply voltage is 1.6 V with a total current drain of 75 mA. Each pixel cell consumes 40  $\mu$ W. The digital supply is 2 V at 40 mA, so the total power dissipation of the pixel IC is 200 mW.

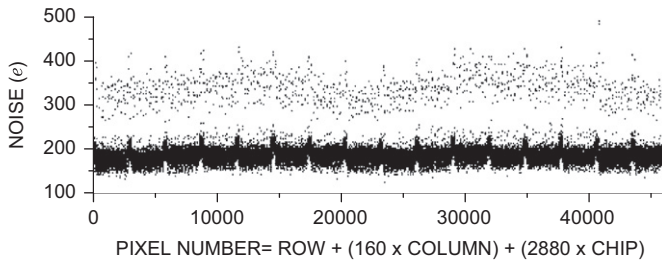
A 40 MHz differential time stamp bus routes timing signals to all pixels. This bus is 8 bits wide and uses a Gray code so that the number of high and low levels remains constant to minimize supply current fluctuations. During signal acquisition each pixel cell records the leading and trailing edge timing. As soon as a trailing edge is recorded, a hit signal (address plus leading and trailing edge timing) is sent to the column periphery (Fig. 44). This operates at transfer rates up to 20 MHz. Differential drivers and receivers are used to minimize cross-talk. At the end of each column pair a content addressable buffer memory with 64

locations is available (one location for each five pixels). Upon receipt of a Level 1 trigger, the buffers are checked for valid events and hits from rejected crossings are cleared. A readout sequencer stores up to 16 events pending readout. Data are “pushed” off the front-end chip to the readout chip without handshaking. All column pairs operate independently and in parallel. The readout operations incur no deadtime, so the overall rate is limited by the buffer capacity [78].

All control data are stored in a 231-bit control register with full readback capability. Configured as a shift register with a shadow latch it utilizes triple redundancy for single event upset tolerance, as it holds critical configuration data. In addition to a broadcast mode, each chip can be uniquely addressed; its identity is controlled by external wire bonds. Intermediate metal layers and special layout limit cross-talk from the digital signals to the sensors and input nodes [77]. Fig. 45 shows the measured noise.



**Fig. 44.** Pixel columns are arranged as pairs of “back-to-back” cells, so the digital control and readout lines are not adjacent to the input pads. When a pixel is struck it sends its address and time stamp to its respective column buffer, configured as a content addressable memory. A trigger selects valid time stamps, which are then sent sequentially to the readout buffer.



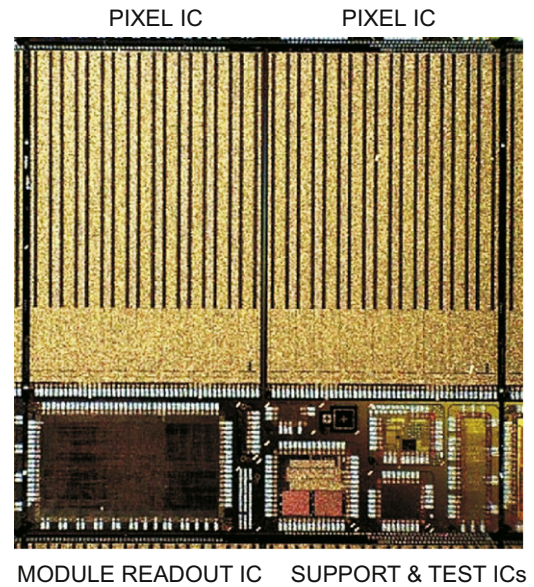
**Fig. 45.** Noise distribution of an ATLAS pixel module. Three groups are visible: the nominal pixels, the extended pixels that bridge columns between ICs (visible as spikes every 2880 pixels), and ganged pixels to bridge rows between ICs. The distributions of pixel noise levels are Gaussian, with  $Q_n \pm \sigma = (184 \pm 11)e$ ,  $(204 \pm 13)e$ , and  $(336 \pm 31)e$ . This module has seven “bad” pixels. (Data courtesy of M. Garcia-Sciveres.)

Ref. [79] summarizes design and layout considerations for pixel readout ICs.

Fig. 46 shows a reticle containing two pixel ICs, a readout controller, and support and test devices. On the pixel ICs the upper 75% are the pixel cells, whereas the lower 25% are readout logic and output drivers. Higher density processes would reduce this area and also allow smaller pixel cells. The 0.25  $\mu\text{m}$  process used in ATLAS production would allow half the pixel size, but the sensor design was determined for an earlier design. Die size is  $7.2 \times 10.8 \text{ mm}^2$ . The edges of the die may not extend beyond 100  $\mu\text{m}$  from the active area. External contacts are by 30 wire bonds, with  $100 \times 200 \mu\text{m}^2$  pads. Diagnostics are provided on 17 additional bond pads.

ICs are thinned to 180  $\mu\text{m}$ . Starting thickness of the wafers is  $\sim 500 \mu\text{m}$ . The chip boundaries are grooved to a depth of  $\sim 200 \mu\text{m}$  and then the wafer is ground to a thickness of  $\sim 180 \mu\text{m}$ , releasing the chips with very smooth edges. Material in a module is about  $0.8\% X_0$ . Including the support structure and services, the material adds up to  $7.5\% X_0$  at normal incidence, increasing to about  $35\% X_0$  at  $\eta = 2$ .

Other pixel designs incorporate many of the techniques described above. CMS also uses a bump-bonded pixel structure, but the pixel size is  $100 \mu\text{m} \times 150 \mu\text{m}$  after migrating to a 0.25  $\mu\text{m}$  process for the readout IC [43]. A column-based readout



**Fig. 46.** A reticle of the pixel IC wafer, showing how multiple ICs are accommodated in one reticle. Reticles are copied by a step-and-repeat process to fill the entire 200 mm wafer. Two pixel ICs are at the top, with the readout controller and test/support ICs at the bottom. The pixel IC is  $7.3 \times 10.9 \text{ mm}^2$  and contains 2880 pixels. (Die photo courtesy of K. Einsweiler.)

is also used, but with a fully analog readout [80]. ALICE uses  $50 \mu\text{m} \times 425 \mu\text{m}$  pixels with a binary readout. A fast OR output can be used in the Level 0 trigger. The proposed BTeV experiment took the readout a step beyond all existing silicon detector systems by reading out at the full beam crossing rate of 7.6 MHz [81,82].

Design and construction of these large-scale pixel systems pose formidable challenges. The ATLAS pixel IC contains nearly four million transistors, so simulation and design verification are crucial. As in all large-scale semiconductor detector systems, electro-mechanical integration – combining sensors, electronics, cabling, cooling, and mechanical support systems – is a major part of the project. The complexity of integrating these systems is usually not appreciated by those who haven’t done it. Furthermore, these

systems are chronically underfunded, as funding agencies, reviewers, and project managers tend to think that these developments merely follow recipes, rather than applying a wide range of physics, so they underestimate the required effort.

## 10. Conclusions

The LHC has imposed unprecedented requirements on new detectors, both because of high flux, i.e. high event rates and radiation damage, and large scale, which requires reliability, complex infrastructure, and worldwide collaborations, even for subdetectors and their components. All detectors are now fully functional, which illustrates that “bottom up” management, i.e. making decisions based on open discussions and verified results, is highly successful. It also illustrates that there are multiple solutions to the same general requirements. This success also built on many prior developments in a wide range of applications that applied novel techniques, which also makes a strong case for targeted generic R&D, rather than constraining by solely project-based funding. The potential sLHC upgrade is raising even more requirements, so ongoing support for new developments is still essential, especially as many of these developments will also encourage other applications.

## Acknowledgment

This work was partially supported by the Director, Office of Science, Office of High Energy and Nuclear Physics of the US Department of Energy under contract no. DE-AC02-05CH11231.

## References

- [1] H. Spieler, Semiconductor Detector Systems, Oxford University Press, Oxford, 2005.
- [2] K. Abe, et al., Nucl. Instr. and Meth. A 400 (1997) 287.
- [3] P. Denes, et al., A fast, direct X-ray detection charge-coupled device, Rev. Sci. Instrum. 80 (2009) 083302.
- [4] S. Worm, et al., Nucl. Instr. and Meth. A 582 (2007) 839.
- [5] S.E. Holland, et al., SPIE J. 6276 (2006) 62760B-1.
- [6] R.H. Richter, et al., Nucl. Instr. and Meth. A 511 (2003).
- [7] R. Turchetta, et al., Nucl. Instr. and Meth. A 501 (2003) 251.
- [8] C. Kisielowski, et al., Microsc. Microanal. 14 (2008) 469.
- [9] Y. Arai, et al., Nucl. Instr. and Meth. A 579 (2007) 701.
- [10] Y. Ikegami, et al., Nucl. Instr. and Meth. A 579 (2007) 706.
- [11] R. Lipton, Nucl. Instr. and Meth. A 579 (2007) 690.
- [12] ATLAS TDR, ATLAS inner detector technical design report, CERN/LHCC/97-16 and 97-17, ISBN 92-9083-102-2 and 92-9083-103-0, available at <http://atlas.web.cern.ch/Atlas/interal/tdr.html>, 1997.
- [13] B. Ashmanskas, et al., Nucl. Instr. and Meth. A 518 (2004) 532; originally proposed by M. Dell’Orso, L. Ristori, Nucl. Instr. and Meth. A 78 (1989) 436.
- [14] J. Adelman, et al., Nucl. Instr. and Meth. A 581 (2007) 473.
- [15] A. Ferrari, et al., Radiation calculations for the ATLAS detector and experimental hall, <http://cdsweb.cern.ch/record/300336/files/est-96-001.pdf>.
- [16] ATLAS Collaboration, 2008 JINST 3 S08003.
- [17] Y. Unno, et al., Nucl. Instr. and Meth. A 511 (2003).
- [18] CMS Collaboration, J. Instrum. 3 (2008) S08004.
- [19] T. Ferguson, et al., Nucl. Instr. and Meth. A 539 (2005) 386.
- [20] W. Snoeys, et al., Nucl. Instr. and Meth. A 465 (2001) 176.
- [21] K. Wyllie, et al., Front-end pixel chips for tracking in ALICE and particle identification in LHCb, in: Proceedings of the PIXEL 2002 Conference, SLAC Electronic Conference Proceedings, Carmel USA, 2002.
- [22] G. Mazza, et al., IEEE Trans. Nucl. Sci. NS-51 (2004) 1942.
- [23] A. Rivetti, et al., IEEE Trans. Nucl. Sci. NS-48 (2001) 1225.
- [24] A. Rivetti, et al., Nucl. Instr. and Meth. A 541 (2005) 267.
- [25] T. Kohriki et al. (1996). NS43/3 (1996) 1200.
- [26] H. Spieler, Nucl. Instr. and Meth. A 636 (2011) 149.
- [27] LHCb Collaboration, J. Instrum. 3 (2008) S08005.
- [28] S. Lochner, M. Schmelling, The Beetle Reference Manual, Note LHCb-2005-105, <http://cdsweb.cern.ch/record/1000429>.
- [29] P. Collins, LHCb Upgrade – A Vertex Detector for High Luminosity, in: Seventh International “Hiroshima” Symposium on the Development and Application of Semiconductor Tracking Detectors, Hiroshima, Japan, September 19–21, 2009.
- [30] ALICE Collaboration, J. Instrum. 3 (2008) S08002.
- [31] W. Snoeys, et al., Nucl. Instr. and Meth. A 439 (2000) 349.
- [32] F. Campabadal, et al., Nucl. Instr. and Meth. A 552 (2005) 292.
- [33] M. Ullan, et al., Nucl. Instr. and Meth. A 604 (2009) 668.
- [34] N.J. Buchanan, et al., J. Instrum. 3 (2008) P09003.
- [35] N.J. Buchanan, et al., J. Instrum. 3 (2008) P03004.
- [36] Third RD48 Status Report, CERN/LHCC 2000-0089, 1999.
- [37] Y. Unno, et al., Nucl. Instr. and Meth. A 383 (1996) 211.
- [38] F. Albiol, et al., Nucl. Instr. and Meth. A 403 (1998) 247.
- [39] S. Gadomski, et al., Nucl. Instr. and Meth. A 320 (1992) 217.
- [40] N. Bingenfors, et al., Nucl. Instr. and Meth. A 326 (1993) 112.
- [41] M. Raymond, et al., The CMS tracker APV25 0.25  $\mu\text{m}$  CMOS readout chip. in: Sixth Workshop on Electronics for LHC Experiments, Krakow, September 2000.
- [42] M.J. French, et al., Nucl. Instr. and Meth. A 466 (2001) 359.
- [43] W. Erdmann, Nucl. Instr. and Meth. A 518 (2004) 324.
- [44] F. Anghinolfi, et al., IEEE Trans. Nucl. Sci. NS-49 (3) (2002) 1080.
- [45] T. Kondo, et al., Nucl. Instr. and Meth. A 485 (2002) 27.
- [46] T. Kohriki, et al., IEEE Trans. Nucl. Sci. NS-49 (6) (2002) 1378.
- [47] G.F. Moorhead, Nucl. Instr. and Meth. A 485 (2002) 43.
- [48] L. Feld, Nucl. Instr. and Meth. A 511 (2003) 183.
- [49] R. Nisius, Nucl. Instr. and Meth. A 530 (2004) 44.
- [50] P. Ferrari, Nucl. Instr. and Meth. A 530 (2004) 38.
- [51] F. Campabadal, et al., Nucl. Instr. and Meth. A 538 (2005) 384.
- [52] W. Dabrowski, et al., IEEE Trans. Nucl. Sci. NS-47 (2000) 1843.
- [53] M. Krammer, Nucl. Instr. and Meth. A 511 (2003) 136.
- [54] S. Braibant, et al., Nucl. Instr. and Meth. A 485 (1997) 343.
- [55] T. Bergauer, Nucl. Instr. and Meth. A 518 (2004) 317.
- [56] M. Krammer, Nucl. Instr. and Meth. A 531 (2004) 238.
- [57] A. Macchiolo, Nucl. Instr. and Meth. A 530 (2004) 54.
- [58] M. Turala, Nucl. Instr. and Meth. A 466 (2001) 243.
- [59] M. Biasini, Nucl. Instr. and Meth. A 530 (2004) 17.
- [60] S. Schnetzer, Nucl. Instr. and Meth. A 501 (2003) 100.
- [61] P. Kuijter, Nucl. Instr. and Meth. A 530 (2004) 28.
- [62] G. Stefanini, Nucl. Instr. and Meth. A 530 (2004) 77.
- [63] K.H. Becks, et al., Nucl. Instr. and Meth. A 386 (1997) 11.
- [64] E.H.M. Heijne, et al., IEEE Trans. Nucl. Sci. NS-42 (4) (1995) 413.
- [65] E.H.M. Heijne, Nucl. Instr. and Meth. A 465 (2001) 1.
- [66] H. Spieler, Integrated microsystems as a driving force in modern detector designs. in: M. Budinich, E. Castelli, A. Colavita (Eds.), International Conference on the Impact of Digital Microelectronics and Microprocessors on Particle Physics, World Scientific, Singapore, 1988, pp. 228–238, ISBN 9971-50-742-0.
- [67] O. Barkan, et al., Development of a customized SSC pixel detector readout for vertex tracking, in: T. Dombek, V. Kelley, G. Yost (Eds.), Proceedings of the Symposium on Detector Research and Development for the Superconducting Super Collider, World Scientific, Singapore, October 15–18, 1990.
- [68] G. Kramer, et al., Development of pixel detectors for SSC vertex tracking. Supercollider 3, in: J. Nonte (Ed.), Proceedings of the 3rd Annual International Industrial Symposium on the Super Collider, Atlanta, Georgia, Plenum Press, New York, March 13–15, 1991, pp. 953–964 ISBN: 0-3064-4037-7, QC787.P7.I57 1991.
- [69] C. Gemme, Nucl. Instr. and Meth. A 501 (2003) 87.
- [70] L. Rossi, Nucl. Instr. and Meth. A 501 (2003) 239.
- [71] J. John, et al., Nucl. Instr. and Meth. A 531 (2004) 202.
- [72] R. Wunstorff, Nucl. Instr. and Meth. A 466 (2001) 327.
- [73] I. Gorelov, et al., Nucl. Instr. and Meth. A 489 (2002) 202.
- [74] R.H. Richter, et al., Nucl. Instr. and Meth. A 377 (1996) 412.
- [75] I. Kipnis, et al., IEEE Trans. Nucl. Sci. NS-44 (3) (1997) 289.
- [76] K. Einsweiler, private communication, 2004.
- [77] L. Blanquart, et al., IEEE Trans. Nucl. Sci. NS-51 (4) (2004) 1358.
- [78] E. Mandelli, et al., IEEE Trans. Nucl. Sci. NS-49 (4) (2002) 1774.
- [79] P. Fischer, Nucl. Instr. and Meth. A 501 (2003) 175.
- [80] M. Barbero, et al., Nucl. Instr. and Meth. A 517 (2004) 349.
- [81] S. Kwan, et al., Nucl. Instr. and Meth. A 511 (2003) 48.
- [82] M. Wang, Nucl. Instr. and Meth. A 501 (2003) 214.